

Patent rules should include a defence against monopolies

Although resentment persists against the patenting of genetic knowledge, the breadth of such patenting is of more urgent concern. The time is right for serious political scrutiny of the issue at a global level.

Few issues are likely to receive as much attention at the World Conference on Science to be held in Budapest later this month as access to scientific knowledge. One aspect likely to be high on the agenda will be that of intellectual property rights over scientific knowledge in general, and genetic knowledge in particular. Some will no doubt seek to use the conference to question again whether it is morally acceptable to claim patents over genetic knowledge at all, on the basis that such data should be considered as a product of nature. But the debate should not be diverted from the more immediate question of whether the breadth and nature of some patents allow their owners too extensive control over this key area of technology.

The issue was quite properly highlighted last month in a report on genetically modified crops produced by the Nuffield Council on Bioethics. Although overshadowed by the council's louder message — the “moral imperative” to assure the development of such crops for use in developing countries — the report also warned against the danger that an entire crop might become controlled by a single company. It called on international plant and patent offices “to avoid granting broad patents that could lead to monopoly suppliers”.

Few would disagree with this sentiment. One reason for the vigorous backlash against the growth of such crops in Britain, for example, is that concerns about the nature of the technology itself have been magnified by fears that it is being introduced into the market-place largely at the behest of a single US corporation, Monsanto. Such worries find an easy resonance among protest and opposition groups in Third World countries. Bearing in mind strong public concern over monopoly control of food production, it is important that the patent system should not be seen as an agent of such monopolization.

Can excesses be avoided through vigilant application of the

current patent system, or is broader political action needed? There is some evidence that current constraints are sufficient, particularly as broad patent claims are highly vulnerable to challenges that the technologies they describe are not truly inventive. This, for example, is the factor that last week allowed critics to prevail against a patent issued by the European Patent Office to a US biotechnology company on the use of stored stem cells from umbilical-cord blood (see page 626).

But there are also dimensions that require political scrutiny, particularly when there are ethical considerations at stake which patent offices are ill-equipped to deal with. One is the question that has been raised by biotechnology critics in the United States of where to draw a dividing line when patenting organisms that contain human cells (see page 626). The US Patent Office itself appears to admit that US legislation is less than clear on this issue. Firm guidelines, ideally resulting from a vigorous political discussion of the consequences, could help to clarify an otherwise muddy area.

This is where the Budapest conference could play a useful role. It would be foolhardy to ask the conference to register broad disapproval of the patenting in principle of sensitive areas of scientific knowledge: too much money is at stake — and too much has already been patented — for such a move to make much sense. But a focused debate on the implications of the breadth of patents currently being issued, the ambiguities and moral uncertainties in current patent legislation, and the political opportunities for addressing both issues (for example, through revisions to the rules of the World Trade Organization) would be welcome. It is to address just such issues that the conference has been organized. Hopefully, delegates will grasp the opportunity presented to them with the commitment and seriousness it requires. □

Boost US infrastructure

Congress should seize a welcome opportunity to strengthen dilapidated university laboratories.

Surprising though it may seem to those in less well-endowed countries, scientists in the United States have good reason to complain of ageing and overcrowded laboratories, obsolescent instruments and daunting costs in maintaining up-to-date specialized facilities. In a biennial survey conducted across the disciplines last year by the National Science Foundation, fully two-thirds of institutions reported overcrowding. And while they said that nearly one-quarter of their research space needed major renovation or replacement, they also said they had put off over \$11 billion in such work for lack of funds.

Given a major decline in defence funding in universities, for example, some suffering in the physical sciences, though regrettable, is not surprising. But in the face of these numbers, and in this era of \$2-billion jumps in the budget of the National Institutes of Health, it is startling to consider that this agency's main programme for funding such construction is spending just \$30 million this year. Similarly,

a shared-instrumentation programme delivers only \$35 million for costly instruments not covered by individual research grants. To be fair, the \$15.6-billion agency also directs some \$3 billion to universities in indirect research costs, some of which can be drawn on for construction and instrumentation. But, especially for smaller and poorer schools, these funds hardly meet the sizeable bill for getting projects off the ground and instruments bought.

Senator Tom Harkin (Democrat, Iowa) has the right idea in a bill that would designate \$750 million for the National Institutes of Health to spend on construction and renovation at its grantee institutions in the next two years (see page 621). Individual scientists may worry that this would eat into their piece of a finite pie. But that would be short-sighted in a golden age in which the pie is growing considerably every year. For the promise of biomedical research to be realized, its underpinnings have to be maintained. Congress should back Harkin's proposal, while not losing sight of the needs of other sciences too. □



UK to help fund US laser fusion research

AP/BEN MARGOT

[LIVERMORE, CALIFORNIA] Britain's Ministry of Defence has announced a significant scientific and financial commitment to the US National Ignition Facility (NIF), the world's largest laser, being built at the Lawrence Livermore National Laboratory.

An initial investment of "tens of millions of US dollars" will fund research by British scientists to allow more rapid repeat firings of the 192 lasers in the NIF's target chamber.

Within 18 months, the ministry will complete a feasibility study of the United Kingdom building a second laser target chamber at the Livermore Laboratory. This could increase UK investment in the project to more than \$100 million.

The announcement of the British commitment to the NIF — the United States' largest single scientific project, costing \$1.2 billion — was made by Graham Jordan, acting chief scientific adviser for the UK defence ministry, at a ceremony last Friday (11 June) marking the unveiling of the NIF's target chamber (see *Nature* 399, 515; 1999).

British scientists will soon join the NIF project at the Livermore Laboratory, said Jordan, although he declined to say how many. Jordan said that US progress on the "massive project" was "very impressive".

Britain's investment will allow its scientists to perform experiments at the NIF, which is causing excitement over its potential contributions to nuclear stockpile stewardship, fusion energy development and



On target: the first laser target chamber at the National Ignition Facility was christened last week.

research in astrophysics, hydrodynamics and high-energy physics.

The UK team will join French scientists, who are already working at the Livermore Laboratory on NIF development. France is constructing its own ignition facility, the Laser MegaJoule, near Bordeaux.

The French facility is being built more slowly than the NIF, as French scientists seek to learn from the US project. Jacques Bouchard, director of military applications for the French Atomic Energy Commission, said at last week's ceremony that French scientists "were more conservative" than the US team on the number of lasers needed for the target chamber. Construction of the NIF is likely to be finished in 2003, with the French facility being completed five years later.

When fired, the 192 lasers will concentrate on a capsule of deuterium and tritium, the two heavy isotopes of hydrogen, forcing them to combine through compression and heating them to produce ignition and self-sustained fusion burn.

Bruce Tarter, director of the Livermore Laboratory, said that confidence was growing over the creation of "a thermonuclear reaction" that could be used for nuclear stockpile stewardship and research.

US Department of Energy (DoE) officials said last week that the NIF project was on schedule and on budget, with \$700 million spent so far on the containment building and the 150-ton laser target chamber it will house.

If Britain goes ahead with the second target chamber, it would be adjacent to the current chamber, allowing scientists to increase significantly the rate of experiments.

DoE secretary Bill Richardson dedicated the NIF target chamber at the ceremony, and said: "This is a landmark toward achieving ignition. NIF will help maintain the safety and reliability of America's nuclear deterrent and expand the frontiers of science."

Richardson said that the greatest value of the British and French contributions to the NIF was "not the sharing of cost, but the sharing of minds".

Jordan said later that participation in NIF will permit the United Kingdom to engage in nuclear stockpile stewardship using the most advanced scientific techniques. **Rex Dalton**

DFG urges Germany to boost its spending on genome research

[MUNICH] Deutsche Forschungsgemeinschaft (DFG), Germany's university granting agency, has urged the German government to provide an extra DM1 billion (US\$530 million) for genome research over the next five years, considerably more than doubling the current level of expenditure.

In a statement published last week, the DFG also proposes that, in order to catch up with other countries such as Britain and France, Germany should create a "national genome initiative" led by the federal science ministry and the rapidly growing biotechnology industry.

It suggests that extra money allocated to bioinformatics be used to set up courses at universities to meet the increasing demand for trained personnel. The DFG also urges the creation of a national committee, headed alternately by representatives of the various areas of genomics, to coordinate the funding of genomic research in Germany.

Despite Germany's "late start" in basic genomics research, the statement, *Perspectives on Genome Research*, says that the scientific quality of such research is

excellent, identifying work on model organisms, microbiology and plant sciences as being particularly strong.

But it points out that, although public funding for such research in Germany has risen, it is still a tenth of that in the United States, where the economic and scientific significance of genome technology was identified much earlier. Fresh money is needed at a time when genomic research is about to enter its next phase, identifying the relevance of genes to the origin of diseases.

At present, public research grants for projects in genomics research in Germany total about DM70 million a year. Of this, DM40 million is reserved for human genome research, with DM5 million for plant genome research, DM7 million for research on microorganisms and DM20 million for technology development.

The DFG suggests that an additional DM400 million be made available for human-genome research and research on model organisms over the next five years. Research on microorganisms and plant sciences should receive DM200 million each,

and bioinformatics and technology development DM100 million each.

Such large sums are needed because international competition is becoming fiercer, says the DFG, pointing out that the French government plans to invest an additional US\$330 million in genome research over the next three years (see *Nature* 399, 185; 1999).

Ernst-Ludwig Winnacker, president of the DFG, points out that Germany has a lot of catching up to do in fields such as functional genomics and bioinformatics. "Now that we have collected so much data we want to know what it is good for," he says.

It is unclear how the DFG's demands could be financed. Public spending in Germany is to be reduced next year by DM30 billion, according to the finance ministry, and the federal budget for science is unlikely to be excluded from cuts. But science minister Edelgard Bulmahn is taking the DFG's proposals seriously, and a task force has been set up by the science ministry to translate them into new funding priorities. **Quirin Schiermeier**

Markl warns Germany not to reduce budget increase

HELLA WOLFF-SEYBOLD

[DORTMUND, GERMANY] Hubert Markl, the president of the Max Planck Society (MPS), has warned Germany's federal and regional governments to expect fierce protests if they try to back down on a promise to increase the organization's budget for 2000 by five per cent.

Finance ministers of the German *Länder* (states) recently proposed that this increase — which would cost DM95 million (US\$51 million) next year — should be reduced to only two per cent, or DM38 million (see *Nature* 399, 508; 1999). They say this is due to the continuing problems in the German economy and the need to cut back on public spending.

But Markl told the annual assembly of the MPS in Dortmund last week that, if such a cut were to be made, the extra money in next year's science budget would be totally absorbed by salary increases, and that "in real terms our budget would shrink".

He pointed out that Rita Colwell, the director of the US National Science Foundation, had said that her agency was expecting an increase of up to 25 per cent next year, and he urged Chancellor Gerhard Schröder and Edelgard Bulmahn, the federal science minister, not to break their "plain promises" — a reference to a pre-election pledge to maintain the funding increases promised by the previous government. Otherwise, he warned, the competitiveness of cutting-edge science in Germany would be seriously weakened.

Markl also warned that if the finance ministers' proposals were accepted, institutes in east Germany would suffer most. Since 1991, annual budget increases have allowed the MPS to set up 19 institutes in the five new *Länder*, and most of the scientific directors for them have now been appointed.

"If we have to back out just before regular scientific work starts, it would end up in very hard and embarrassing cancellations of appointments," said Markl.

Markl argued that the new financial constraints on basic research are the result of misguided political decisions and priorities. He pointed out, for example, that Germany's 40 per cent contribution to the European component of the International Space Station had been agreed "against the advice of many experts".

"Science did not want a project that would eat up more than \$100 million," Markl said. He argued that Germany's "political decision" to contribute to the space station should not be financed by cuts in the basic science budgets.

In his speech, Markl also emphasized



Markl: east Germany would be hit hardest.

that the MPS was keen to implement changes recommended last month by an external evaluation committee (see *Nature* 399, 395–396; 1999), as he conceded that the society had occasionally neglected modern research needs through "self-deception" and "insufficient public relations".

At the same time, he warned society members not to act too hastily, and he said that all recommendations should be carefully considered over the next six months. Wolf Singer, director of the Frankfurt-based Institute for Brain Research, has been chosen to head a committee that will draft proposals for reform by the autumn.

The MPS has already taken up one recommendation: to improve career opportunities for women. At present, there are only five women among the 241 directors holding top research (C4) positions. Over the next three years, 15 additional C4 positions are to be created exclusively for women.

"We will start to actively look for female candidates, even though we are aware of possible legal conflicts over discrimination against men," said Markl.

This programme was approved by the MPS's senate provided that the five per cent increase in the budget — which the evaluation committee also recommended — was not cut.

Markl has sent protest letters to the prime ministers of the 16 *Länder* — which will make the final decision on the MPS budget — and to the Bund-Länder Kommission, the body that coordinates regional and federal research policies.

Jörn Brand, head of the Bund-Länder Kommission's research funding department, says the finance ministers' stance will be difficult to shift. "It is no longer justifiable that the MPS should be treated so much better than the national research centres or blue-chip institutes," he says.

But Brand is confident that the federal government will recommend a compromise proposal increasing the budget by around three per cent. The MPS should then be able "to live with what is still an increase", he says.

Quirin Schiermeier

Referee quits journal over price rise as library faces cutbacks

[PARIS] A physicist at Florida State University has resigned as an unpaid referee for a nuclear physics journal published by Elsevier Science in protest at inflation in the price of the company's journals. His home institution supports his decision.

Mark Riley says he was "upset" at cutbacks in journal subscriptions at his university. Its library, like many worldwide, has faced annual inflation in journal prices of around 20 per cent over the past decade. As a result it has cancelled some 2,000 titles and plans to cut a further \$800,000 worth of journals over the next three years.

In a letter of resignation to Jeanette Bakker, executive editor of *Nuclear Physics A* — which costs \$7,234 a year — Riley wrote: "I am dismayed by the pricing and inflation policies of Elsevier and the significant part they have played, and are playing, in the present journal budget crisis." He adds: "While I respect the quality of *Nuclear Physics A* as a scientific journal, I feel honour bound to minimize my involvement with Elsevier Publishing at this time."

Riley says that other referees "should think about what they are doing in refereeing for expensive journals, and remember that the serials crisis is affecting their home institutions".

Charles Miller, dean of libraries at the university, applauds Riley's decision. Miller says that journal subscriptions account for 75 per cent of the libraries' spending, and that around 25 per cent of this is on Elsevier Science journals.

"We want to go all-electronic, delivering journals to the desktop and phasing out print altogether," says Miller. But he says this shift is hampered by reluctance on the part of Elsevier Science to dissociate paper and electronic access in its pricing policies.

A new contract with the company has brought down prices, and Miller speaks of "a willingness on Elsevier's part to bring them down further". The university is demanding that Elsevier provide electronic-only access to all its journals at a 30–40 per cent discount on current prices. So far Elsevier has offered 10 per cent, Miller says.

He warns that the university may cancel expensive journals even if these are in heavy demand by users. "They don't believe we will cancel, but we will."

Elsevier Science was unavailable for comment. But Geert Noorman, managing director of its Life Sciences Division, recently told *Nature* that web publishing is putting pressure on publishers: "If the web causes us to have to agree to lower profit margins, then so be it."

Declan Butler

Japan may lift industry–academy barrier

[TOKYO] Japan has announced plans to set up an inter-ministerial committee to explore ways to make university researchers more free to collaborate with industry.

The committee, whose creation was announced by prime minister Keizo Obuchi last week, will decide by the autumn whether to recommend a relaxation in the current civil service law. This forbids employees of national universities from taking part in profit-making activities.

Obuchi has set up the committee in response to controversy over the appointment by the electronics company Sony of an economics professor from Hitotsubashi University, Iwao Nakatani, to its board of directors (see *Nature* 398, 643; 1999).

The Ministry of Education, Science, Sports and Culture (Monbusho), had intended to exempt Nakatani from a law prohibiting academics from holding external posts. But this move was rejected by the National Personnel Authority, the government's central personnel agency, leading Nakatani to resign from his university post.

The plan to revise the current civil service law is part of the government's latest job-creation package. The ¥500 billion (US\$4.2 billion) package, approved by parliament last week, aims to create 700,000 jobs by supporting fast-growing industries — such as

biotechnology and information sciences — and promoting the commercial application of university research.

The package includes measures to allow technology licensing organizations (TLOs) — which advise on the commercial exploitation of university research — to hold patent rights to inventions made at universities.

Under the current law, rights to patents produced from national universities belong to the national treasury, and cannot be assigned to private companies without complex legal procedures.

Both industry and universities hope that the plan will improve links between industry and academia, and encourage university researchers to set up venture businesses.

The Ministry of International Trade and Industry (MITI) runs a programme with Monbusho to set up TLOs and venture businesses at universities. It supports the involvement of university scientists in business.

"If university researchers were allowed to take up positions in private companies, this could stimulate entrepreneurial spirit within Japanese universities," says a MITI official.

Current restrictions mean that ventures set up by university researchers have to recruit staff from outside the university, or appoint postgraduate students as managers.

Examples are Tokyo University's Centre

for Advanced Science and Technology Incubation (see *Nature* 396, 9; 1998) and Synthesis, a company set up by professors from Osaka and Kyoto universities.

"It seems silly that we have to make so many compromises and overcome numerous restrictions in order to run a business," says Kozo Murakami, professor in engineering at Osaka University and a founder member of Synthesis, which designs and develops large-scale integrated circuits.

The company is managed by postgraduate students, while the professors are only allowed to give technical guidance. "I hope we would be able to take a more active role in management if the restrictions were to be eased," he says.

Although some universities still disapprove of academic staff becoming involved in business, the private sector has welcomed the government's move to encourage collaboration with university researchers.

"Companies would benefit a lot from gaining a voice from the academic community," says Takuji Nakamura, of the Kansai Technology Licensing Organization, which helps draw up contracts for collaborations between universities and industrial laboratories. "But we must first remove the legal barrier that restricts effective communication between the two worlds." **Asako Saegusa**

South African government seeks reassurances on AIDS initiative

[CAPE TOWN] South Africa's Minister of Health, Nkosazana Zuma, has confirmed she has not rejected an offer by the US pharmaceutical company Bristol-Myers Squibb (BMS) to donate \$100 million over the next five years towards fighting AIDS in southern Africa (see *Nature* 399, 96; 1999).

But, speaking on South African radio, Zuma said that the offer must be modified to conform with the government's healthcare policy before it is endorsed.

The initiative, dubbed 'Secure the Future', is targeted at South Africa, Botswana, Lesotho, Namibia and Swaziland. On its announcement last month, the South African Department of Health said it wished to be "an active partner in this effort".

According to Jon Weisberg, a spokesman for BMS, one important feature of the initiative is the establishment of a 'virtual institute', funded by BMS. This will assess research proposals from southern African universities and health research bodies.

But the department also said that the government's endorsement was "conditional on the development of a mutually acceptable implementation plan for the South African component".

Zuma outlined four areas of concern at a



South Africa's health minister wants AIDS care to prioritize orphans — shown here in Natal.

meeting with BMS before its announcement: sending South African doctors to the United States for training in HIV treatment — including clinical trials of BMS products (which Zuma rejected as irrelevant to the country's needs); financing clinical trials in South Africa; that only preclinical studies approved by the South African Medicines Research Council were acceptable; and that funding to community outreach programmes should focus on giving health care at home and caring for orphans.

Ian Roberts, special adviser to Zuma,

said that Zuma told BMS that South Africa would not endorse their programme until it could view and approve the alterations to the initiative. But BMS left the plan unchanged when it was launched, and two weeks later Roberts issued a statement in which he emphasized that the government had not endorsed the initiative (see *Nature* 399, 288; 1999).

But according to Mark Ahn, senior director, operations and planning (international), at BMS, Zuma had, before the meeting, appointed the heads of the Medical Research Council of South Africa and the department's HIV/AIDS/STD Directorate to the advisory board that will select the research projects to be funded.

Such studies are to be proposed by African medical institutions, and not the company. The board will also select the fellows who will participate in the health education exchange programmes.

"We are still awaiting a response from BMS on whether they are prepared to tailor their proposals to South Africa's needs," says Vincent Hlongwane, Zuma's spokesman. It is hoped that the position is clarified soon, as the first board meeting of the initiative is scheduled for 24 June. **Michael Cherry**

Congress limits moratorium on lab visits

[WASHINGTON] The US House of Representatives last week approved a 60-day moratorium on visits to US nuclear weapons laboratories by visitors from 'sensitive countries'. This forestalls the enactment of a more radical two-year moratorium.

The moratorium was approved as part of a package of 26 measures aimed at tightening laboratory security in the wake of the Chinese spy scandal. These were passed as part of a bill authorizing spending by the Department of Defense in the year 2000.

The 428-0 vote on the package, authored by congressmen Christopher Cox (Republican, California) and Norm Dicks (Democrat, Washington), virtually assures that the measures will be retained when House and Senate conferees meet to reconcile differences between their respective versions of the bill.

The moratorium will start 30 days after the bill becomes law, and is intended to be in

effect while the Department of Energy (DoE) institutes new vetting procedures for foreign scientists. The Secretary of Energy may waive the moratorium on a case-by-case basis.

The bill applies to foreign visitors to Lawrence Livermore National Laboratory in California, Los Alamos and Sandia National Laboratories in New Mexico, and The Oak Ridge National Laboratories in Tennessee.

A harsher, two-year moratorium on visits by scientists from sensitive countries including China, India and Russia was proposed by congressman Jim Ryun (Republican, Kansas) but defeated on a vote of 159-266.

The votes came two days after the chief science adviser to President Clinton warned that a "xenophobic" Congressional reaction to allegations of Chinese spying could imperil scientific excellence and national security at the three US nuclear weapons laboratories.

In a speech to the US Civilian Research



and Development Foundation, Neal Lane, the director of the White House Office for Science and Technology Policy, said that a Congressional moratorium on lab visits by foreigners "would severely hamper our efforts to control the post-Soviet arsenal of weapons of mass destruction, not only because it would block collaborative activities here, but also because it would immediately lead to curtailment of US access to sites in Russia".

He warned that a moratorium would hurt the labs' participation in international science, calling the proposals "bad for science — and bad for the nation".

Also last week, an advisory committee to DoE secretary Bill Richardson published a report backing international collaborations at DoE laboratories, which it described as "essential to the scientific and technological strength of the United States".

"These types of collaborations ... can be conducted without jeopardizing national security and should be continued," says the report from a working group of the Secretary of Energy Advisory Board, which was mandated in March to examine the department's Foreign Visits and Assignment Program.

Ryun challenges both Lane's criticisms and the findings of the report. "We cannot continue to sacrifice our national security in the name of science," he says. "Many of our most sensitive national secrets have been stolen, while there has been little progress in non-proliferation efforts throughout the world. It is time that Congress addressed the lack of accountability for security in our national nuclear laboratories."

The report to Richardson comes on the heels of a joint statement from the presidents of the National Academy of Sciences, the National Academy of Engineering and the Institute of Medicine, that "inappropriate restrictions on foreign visitors to DoE laboratories could weaken the United States scientifically and prompt retaliation" (see *Nature* 399, 294; 1999).

Meredith Wadman

Energy secretary reassures Asian Americans

[LIVERMORE, CALIFORNIA] Department of Energy secretary Bill Richardson began a morale-boosting tour of national laboratories last week to reassure Asian-American scientists that the recent Chinese spy scandal has not placed them under a broad cloud of suspicion (see above).

In the first such event, Richardson made a speech and took questions at a private meeting last Friday (11 June) with a largely Asian-American audience at the Lawrence Livermore National Laboratory, where he was dedicating the laser target chamber of the National Ignition Facility (see page 622).

Richardson is planning to make similar addresses next week to Asian Americans at the Los Alamos and Sandia National Laboratories in New Mexico.

The Los Alamos meeting may be particularly tense given the international scrutiny of the spy scandal there, in which federal authorities suspect a Taiwanese-born scientist gave some of America's most sensitive nuclear secrets to the Chinese. No criminal

charges have been filed against the scientist, Wen Ho Lee, who was fired in March for security violations.

During the public dedication ceremony for the National Ignition Facility, Richardson told an audience of more than 500 scientists, laboratory employees and their families that the laser project couldn't be accomplished without the assistance of Asian-American scientists.

After the ceremony, Richardson spoke to nearly 400 Asian-American scientists and Livermore Laboratory personnel at a private meeting which officials had attempted to keep secret. Acting under instructions from the Department of Energy in Washington DC, Livermore officials gave false information about the secretary's schedule that day to prevent reporters from knowing of the private discussion.

According to Dorothy S. Ng, however, a civil engineer and structural analyst at Livermore who attended the meeting, representatives of long-established Asian-American groups at the

laboratory presented Richardson with considerable information reflecting their concerns. After the meeting, Ng, who was born in China but is a naturalized US citizen, said she was encouraged by the treatment she and other Asian Americans had received.

"We had a nice dialogue," said Ng, who has worked at the laboratory for 21 years. "We as a group have received a commitment from Richardson and Livermore management to try to correct the negative image. I am impressed that Livermore Laboratory management and the secretary are concerned about us." Congressmen who talk about the loyalty of Asian Americans at national laboratories, Ng said, should be "responsible in their comments and speak of proof and truth".

Livermore officials said that no scientist has quit their laboratory in the wake of the spy scandal. But Richardson is reported to have told the private meeting that one distinguished Asian-American researcher was said to have left the Los Alamos laboratory because of perceived hostility.

Rex Dalton

Patent on umbilical-cord cells rejected in Europe...

[PARIS] Researchers last week won their challenge to a European patent on the use of stored stem cells from umbilical-cord blood, a technology with promising applications in bone marrow transplants and gene therapy. The patent was granted three years ago to the US company Biocyte.

Cord blood stem cells can produce red and white blood cells and platelets, and their transplantation is more effective and cheaper than the conventional practice of taking stem cells from bone marrow donors. Their lower immunogenicity, for example, reduces the risk of rejection by the patient.

But the potential threat of expensive litigation over claims of patent infringement has deterred companies and research groups from exploring new uses for cord blood cells, according to Eliane Gluckmann, a researcher at the Hôpital St Louis in Paris. "That threat has now been lifted," she says.

Gluckmann chairs Eurocord, the European Cord Blood Bank, which involves 14 research teams. It has led the challenge to the patent, along with several other organizations including the environmental group Greenpeace, the US biotechnology company Thermogenesis, and Astra Pharmaceuticals (see *Nature* 382, 99; 1996).

The patent covered "hematopoietic stem and progenitor cells of neonatal and fetal blood, that are cryopreserved, and the therapeutic uses of such cells upon thawing," and gave the company broad rights. The groups had opposed it on ethical grounds, arguing that it was wrong to patent human tissues. But they fought the legal battle on the basis of classical patent law, claiming that the technology described in the patent was not new.

Last week, the European Patent Office in Munich upheld their claim. According to an official, the patent "lacked novelty or an inventive step". The patent office was convinced by several papers describing similar techniques that predated the patent. The official adds that a key element was testimony by Pablo Rubenstein, a researcher at the New York Blood Center, that he had used the techniques long before the patent was applied for.

Rubenstein describes the result as a victory for proper attribution of innovation. He argues that, although the patent was defeated on legal grounds, it is also an ethical victory, as it overturns a patent on human tissues.

Neither Biocyte, which is free to appeal, nor its lawyers were available this week for comment. The patent has already been rejected in the United States and Japan. **Declan Butler**

US agency seeks to boost funds for basic research in Russia

[WASHINGTON] The US Civilian Research and Development Foundation (CRDF), set up to aid scientists in the former Soviet Union, may return to its original practice of sponsoring basic, investigator-initiated research if a hefty increase in funds proposed by the Clinton administration is approved by Congress.

The CRDF, which awards grants averaging \$50,000 for joint research between scientists in the United States and the former Soviet Union, was set up in 1995 under the auspices of the National Science Foundation, using \$5 million in start-up funds from the US Defense Department and \$10 million from philanthropist George Soros (see *Nature* 375, 170; 1995). The foundation's budget is around \$10 million a year.

Principal support has come from the National Science Foundation, the National Institutes of Health and the Department of State. Another 35 or so US agencies have used the CRDF as a conduit for funding cooperative projects with former Soviet Union scientists, according to Gloria Duffy, chief executive of the Commonwealth Club of California, and chair of the foundation's board of directors.

While that has kept the money flowing, it has confined the research agenda to areas of interest to sponsors. According to Duffy, the CRDF wants a separate endowment to enable it to return to its original practice of funding a range of investigator-proposed research, rather than merely "performing a federal agency's task".

At a conference held in Washington last week to highlight results from CRDF collaborations, White House science adviser Neal Lane reiterated the administration's support for increased aid to Russian scientists. President Bill Clinton has proposed that spending should be tripled from \$64 million this year to \$176.5 million in 2000. Annual funding for the CRDF would increase from \$10 million to \$23.5 million, and it would receive \$111 million over five years.

Support for the CRDF is strong in Congress, claims Duffy. But factors such as US-Russian tensions over the war in Kosovo and concerns about security in US government laboratories make it difficult to predict whether law-makers will grant the administration's spending request.

The proposed boost comes at a critical time for former Soviet Union scientists. With the Russian government unable to support its own research establishment, and with Soros pulling back on his commitments (see page 628), Russian scientists have become heavily dependent on US and European aid to continue their work. **Tony Reichhardt**

... as US bid to patent human-animal hybrid fails

[LONDON] The US Patent and Trademark Office has rejected a provocative patent application filed by two prominent biotechnology critics on techniques for combining human and animal cells to create hybrids or chimaera.

The patent was filed by researcher Stuart Newman of New York Medical College, a prominent member of the Council for Responsible Genetics, and Jeremy Rifkin, president of the lobbying group the Foundation on Economic Trends. Both are keen to open up political debate on what can and cannot be patented when human cells are involved (see *Nature* 392, 423; 1998).

In rejecting the application, the patent office lists prominently among its objections the fact that "the claimed invention as a whole

embraces a human being", and that its subject matter therefore lies outside the scope of US patent law.

The examiner who rejected the application admitted that the patenting of human beings was not explicitly ruled out in an earlier judgement, the famous Chakrabarty case in the early 1980s which cleared the way for patents on living organisms – and the growth of the US biotechnology industry.

The examiner lists a variety of reasons for rejecting the application, including the fact that several previous efforts at introducing human cells into animal tissues have been described in the scientific literature. She also argues that, despite the Chakrabarty judgement, in drawing up patent legislation "Congress

did not intend [the US Patent Act] to include the patenting of human beings".

But Pat Coyne, the Washington-based attorney who is handling the patent application, says: "We do not think this is a proper ground to reject the application. We feel that the key issue is: what does it mean to 'embrace' a human being? How does the patent office get the authority to say that, because in its extreme form our claimed invention could do this, it is not patentable?"

Rifkin says: "No parliament in the world is going to be keen to debate how much human genetic information [in a chimaeric organism] makes up a human being. But we want to force them to do it." He and Newman are appealing against the rejection of their application. **David Dickson**

US Congress could double information technology research

[WASHINGTON] US government spending on information technology (IT) would be nearly doubled at six agencies over the next five years under legislation introduced in the House of Representatives last week.

James Sensenbrenner (Republican, Wisconsin), chairman of the House Science Committee, co-sponsored the bill with George Brown (Democrat, California), the committee's senior Democrat, and 23 others.

The Networking and Information Technology Research and Development Act would authorize almost \$4.8 billion worth of IT research between 2000 and 2004 at the agencies in the committee's jurisdiction. They are the National Science Foundation, the space agency NASA, the Department of Energy, the National Institute of Standards and Technology, the National Oceanic and Atmospheric Administration and the Environmental Protection Agency.

Soros cuts off money for Russian science teaching

[MOSCOW] The American philanthropist George Soros is ending his multi-million

dollar programme of support for education in the natural sciences in Russia. In the past five years he has spent \$76 million on grants to school teachers, university professors, and postgraduate and undergraduate students.

Soros had signed an agreement under which the Russian government had agreed to cover half the costs, but Russia failed to find the funds. "Because of the crisis in Russia I prolonged my programme for one year, but I can't do it forever," says Soros. The only exception is \$200 stipends for 260 professors over 70 years of age.

US sounds alarm over smallpox weapon threat

[WASHINGTON] North Korea, Iraq and Russia are likely to be harbouring stores of smallpox virus intended for military use, according to a US government intelligence analysis reported in *The New York Times*. The newspaper reported on 13 June that the government based its conclusion on the analysis of years of data — including the testimony of a former top official in the Soviet Union's germ warfare programme.

The conclusion that the three countries were probably secretly housing the virus is said to have played a key role in President Bill Clinton's decision in April to reverse the US position and retain a store of the virus

housed at the Centers for Disease Control in Atlanta (see *Nature* 398, 741; 1999).

Last month, the World Health Organization pushed back to June 2002 its deadline for deciding whether known existing stores, in Russia and the United States, should be destroyed.

Former Brussels research supremo Fasella dies

[MUNICH] Paolo Fasella, for many years director-general of the European Commission's research directorate, died last week after heart surgery at the age of 68. Following his retirement from Brussels in 1995, Fasella became director of science at the Italian research ministry, and had been appointed to Italy's new strategic research advisory committee, which reports directly to the prime minister.

Research plea to save California's inland sea

[SAN DIEGO] An international panel of scientists has called for research to find new ways to integrate environmental and agricultural needs to maintain the threatened Salton Sea, California's largest inland body of water and an important site for migratory birds.

nature

the international weekly journal of science

needs a Web Editor

Are you seriously fascinated by the Web?

Do you seriously love science?

Would you like to be in charge of the Web presentation of the world's most innovative and, sometimes, publicly controversial science?

Do you have strong editorial experience and flair? Do you have a solid scientific background at least to graduate level?

If your answer is "yes" to all of these, then you could be the person we are looking for. *Nature*, the world's leading international weekly journal of science, needs a Web Editor to take the lead in our development of online services and content.

Our site is already heavily visited by scientists worldwide. The challenges would include the development and launch of new services from the group of *Nature* titles online: *Nature* itself, and six related monthly publications. You would also lead editorial innovation in presenting our content, week by week and archivally, on the Web. And we'd want strong input to strategic



thinking about online science publishing across all disciplines, in a highly competitive and fast-changing environment.

It's a serious job — but stimulating, too.

After all, we want www.nature.com to be THE place for the world's liveliest and most demanding scientists.

This is a senior post, based in our London office, reporting to the Editor of *Nature*. Salary and terms are highly competitive.

Applicants should send a CV, examples of relevant work, and a brief summary (1000-word maximum) of strengths and weaknesses of *Nature's* web site, and their priorities for its development. Mark the envelope "Web Editor", and send to **Dr Philip Campbell, The Editor, Nature, 4 Crinan St, London N1 9XW**, to arrive no later than Monday 12 July. Applications may also be sent by e-mail to webeditor@nature.com.

A new federal law designed to enhance the Salton Sea is spurring discussions of expensive and misguided water projects that will not enhance the bird sanctuary and may damage sensitive ecosystems nearby in Mexico. That is the conclusion of a report just released by the University of California Institute for Mexico and the United States, based in Riverside, California.

Don't pull plug on nuclear power, say UK scientists

[LONDON] Britain's leading scientists and engineers have made a strong plea for the revival of nuclear power. A report from a working group of fellows of the Royal Society and the Royal Academy of Engineering says it is unclear whether renewable energy sources will be able to meet future energy needs. It adds that "it is vital to keep the nuclear option open" to meet energy needs without contributing to climate change.

The report, *Nuclear Energy — The Future Climate*, published on Monday (14 June) calls for the setting up of an intergovernmental research funding agency to investigate sources of electricity generation. The agency's annual budget would be US\$25 billion, funded from national contributions on the basis of gross domestic product or energy consumption.

NZ bid to win rights for apes fails in parliament

[SYDNEY] New Zealand researchers have failed in their bid to persuade parliament to give legal rights to gorillas, chimpanzees, bonobos and orang-utans. This probably spells doom for their efforts to convince the United Nations to make a Declaration of the Rights of Great Apes, based on the genetic proximity of apes to humans (see *Nature* 397, 555; 1999).

The scientists had sought an addition to an animal welfare bill to outlaw the killing of apes, and their use in experiments. But a parliamentary committee rejected a reference to basic rights for apes, and only supported the prevention of experiments on apes.

NASA mission will study origins of the Universe

[BOSTON] The Far Ultraviolet Spectroscopic Explorer (FUSE) is due to be launched from Cape Canaveral next week (23 June) in the hope of gaining an insight about the origins of the Universe and the evolution of galaxies. The project is led and managed by investigators from Johns Hopkins University in Baltimore, Maryland. It is one of the first missions in the US space agency NASA's Origins Program.

Scientists plan to obtain data from FUSE for three years, observing in ultraviolet wavelengths that are inaccessible to other telescopes. The research team expects to learn about conditions in the earliest moments after the Big Bang. The work could "make a leap towards understanding how the primordial chemical elements were created and distributed since the beginning of time".

Inquiry clears ethics of child aggression study

[WASHINGTON] A New York psychiatric institute criticized last year for conducting experiments on aggression in healthy, non-white boys using the banned diet drug fenfluramine was exonerated last week by the US government.

The Office for Protection from Research Risks (the OPRR) at the National Institutes of Health found that the ethics board at the New York State Psychiatric Institute followed government ethics rules in experiments to test whether serotonin, a neurotransmitter, predicts violent behaviour (see *Nature* 392, 747; 1998 & 392, 853; 1998).

But in separate cases the OPRR found that the Mount Sinai School of Medicine and the Research Foundation of the City University of New York breached ethics rules in studies using fenfluramine in children.



The full text of items on this page — and further details and background information about the Unesco/ICSU World Conference on Science, to be held in Budapest from 26 June to 1 July — can be found on *Nature's* website at <http://www.nature.com>

New Zealand and Australia call for wider access to knowledge

[SYDNEY] Australia's delegation to the World Conference on Science in Budapest is expected to urge Unesco to prioritize access to knowledge for developing countries.

The New Zealand delegation will call on Unesco to widen the participation of women and indigenous peoples in science. New Zealand also plans a national conference to be held within ten months of the Budapest meeting to implement its recommendations.

Both delegations will call on Unesco to support Pacific island countries in their efforts to manage their marine resources, slow down climate change and mitigate natural disasters.

The Australian delegation will be headed by John Zillman, director of the Bureau of Meteorology and president of the World Meteorological Organization. The New Zealand group will be led by Paul Ramsay, head of the international affairs branch of the Ministry of Research, Science and Technology. Both delegations will include representatives of indigenous peoples (Aboriginal and Maori) as well as doctoral students.

Delegates hope to initiate an international programme for "enhancing the professional quality of reporting of science, engineering and technology in the mass media". They will seek the agency's support for a scheme similar to the planned International Centre for Science Communication in London.

Full text: <http://helix.nature.com/wcs/a44.html>

Non-governmental groups welcomed as observers

[LONDON] Invitations to attend the World Conference on Science in Budapest later this month have gone out to 150 non-governmental organizations (NGOs). Most of these have gone to organizations of scientists, although Unesco officials say that other groups, such as environmentalist NGOs, are welcome to apply.

The NGOs will only have observer status at the conference. But they may be allowed to address its plenary sessions and to make interventions during debates, say officials.

Representatives of Greenpeace and Friends of the Earth are expected to attend, although there is no environmentalist voice on the main preparatory committee. This was set up to coordinate NGO responses to the conference draft documents. Mauricio Iaccarino, Unesco's assistant director-general for science, said in London this month that "no anti-science or pseudo-science people have been invited [to Budapest]".

The NGO committee is made up of representatives of the Third World Academy of Sciences, the International Council for Social Sciences, the World Federation of Scientific Workers, and the International Council for Philosophy and Human Sciences. It was asked to begin consulting NGOs at the end of May.

One senior member of a science NGO says there are good reasons for restricting attendance from environmental NGOs, even

though these have had a strong presence at meetings of other UN agencies, such as the World Health Organization and the environment conventions. "There will be no mass jamboree. The meeting will be more focused, and we will get things done."

A Unesco spokesman says there is no attempt to exclude environmental groups, adding that he is still processing applications from those who want to attend. He says that science-based organizations are predominant because of their affiliation to Unesco or to ICSU, the International Council for Science, which are jointly organizing the conference. He points out that NGOs are often members of official government delegations.

NGOs have been allocated two half-day sessions during the conference, on 27 and 28 June, to comment on and suggest amendments to the draft *Declaration* and *Framework for Action* documents.

Unesco has also set up a committee to coordinate the NGO community's responses to the draft conference documents, with a deadline for submissions of 20 June, says Andre Jaegle, president of the Paris-based World Federation of Scientific Workers, and a member of the committee.

Conference secretary Howard Moore says: "We hope there will be some controversy, and that it is not just a talking shop for politicians."

Full text: <http://helix.nature.com/wcs/a43.html>

Private research and public access must go together, says Mayor

[BUDAPEST] The most crucial issue facing the World Conference on Science later this month will be how to defend public access to scientific knowledge while benefiting from private investment in research, says Federico Mayor, the director-general of Unesco.

"Science is part of the crucial question of how to ensure that the 'knowledge society' does not become a new form of imperialism, placing developing countries in new chains," says Mayor in an opinion article written for *Nature's* web-based coverage of the conference.

He says that finding solutions to intricate problems will take time. But action can be taken immediately, and issues do not have to be couched in confrontational terms.

"For example, the argument that biotechnology is the key to greater food production has been countered by claims

that it is in fact enhanced democracy that increases food production and distribution," says Mayor. "But we need both; this, surely, has to be the basis of the required relationship between science and society."

Mayor suggests that the two central documents for the Budapest meeting — the *Declaration* and the *Framework for Action* — connect up vision, principles and practice in a "shared dynamic", expressed in every area from forms of funding and training, to strategies for networks and communication.

In innovation and technology foresight, for example, the renewed commitment by governments to fundamental research being sought in Budapest by the scientific community "should ensure that basic science continues to produce the new knowledge without which innovation will ultimately dry up," he says.

New initiatives in the transfer of scientific knowledge, expected to be announced in Budapest, "will attempt to open up global access to the huge store of existing knowledge". And a redefinition of wealth creation to include the massive indirect wealth created by sustainability and preventive measures "must be the driving vision behind science for development".

This coherent, global momentum "requires every single participant at the World Conference on Science and stakeholders in science around the world to join forces and play a role," says Mayor. A universal readiness to get involved in the many initiatives now emerging "can tip the balance and possibly even reverse today's extraordinary concentration of science within a few countries".

Full text: <http://helix.nature.com/wcs/c21.html>

Latin America and the Dracula problem

Sir — In your supplement on *Science in Latin America*, you discuss the relationship between scientists and Mexican industry, but fail to address the question of the development of a scientific culture among the population at large (*Nature* **398** Suppl. 1 April, A7–A9; 1999). This is an important issue, especially as in Latin American countries the health of the research enterprise may depend strongly on its coupling to both education and science popularization among the wider public.

Our societies are rapidly moving in many different cultural directions, but so far science has been largely absent or has played a token role. Science popularization in Mexico is hardly facilitated by the academic establishment. We have excellent science museums, but many institutions and researchers are contemptuous of efforts to promote public understanding of science, an activity which is unfairly considered a flamboyant, second-rate

endeavour. In Mexico a significant portion of a researcher's salary comes from supplements that reflect the number of articles published in refereed journals. Such a policy discourages creative involvement in promoting public understanding of science.

This is a mistake. Like Count Dracula, science should always be on the lookout for fresh blood. Half of our population is aged 15 years or less. To expose young people to an environment in which science is recognized as an essential component of modern culture should be seen as an international responsibility. The intellectual potential destroyed by discouraging a student in a developing country from pursuing a career in science is a loss not only for their country, but for the global scientific community. Such losses could be mitigated by developing multinational efforts to promote science, eagerly but without paternalistic attitudes, with the assistance of international

scientific societies and organizations [see Commentary in this issue, page 633].

This task cannot be merely pedagogical, nor restricted to the utilitarian aspects of scientific research. It is not an easy endeavour, yet we have extraordinary precedents. In colonial Mexico in the late eighteenth century, Jos Ignacio Bartolache, a physician, founded *El Mercurio Volante*, a small newspaper that published advances in medicine, chemistry, physics and astronomy. Bartolache sometimes wrote not only in Spanish, but also in Nahuatl, one of our main Indian languages, so that the benefits of knowledge would be available to a wider segment of the population. A mad, heroic effort. But shouldn't every pioneer have a dose of quixotic blood running in their veins?

Antonio Lazcano

Facultad de Ciencias, Universidad Nacional Autonoma de Mexico, Apdo Postal 70-407, Cd Universitaria, 04510 Mexico D.F., Mexico

Speaking up for our Japanese colleagues

Sir — I have just returned from yet another seminar where the speaker scrupulously named all his colleagues and competitors when referring to their work, except for those from Japan, who were lumped together as "a Japanese group".

As I assume that most of my fellow scientists are not racists, I can only infer that they are too lazy and/or discourteous to bother to remember unfamiliar Japanese names.

If Japanese scientists can learn Western names, why can't we learn theirs?

Kathy Weston

CRC Centre for Cell and Molecular Biology, 237 Fulham Road, London SW3 6JB, UK

Extinction needn't be for ever

Sir — The questions raised by Peter D. Moore¹ in connection with the discovery² of the wild gametophytes of the celebrated British rarity, Killarney fern (*Trichomanes speciosum*) are of great significance. Moore questions the appropriateness of declaring a plant species extinct without being sure that the last individual has been wiped out. We would like to examine these issues in the Indian context, where the appearance

of new taxa, extinction and rediscoveries go hand in hand.

It has been predicted that, in the next few decades, one-third of Indian biodiversity may become extinct or nearly extinct³. At present, the rarity of Indian plants is commonly inferred from herbarium data. If a species has not been collected within the past 50 years, it is considered 'possibly extinct'⁴. This highly biased concept has led to the rediscovery of many 'extinct' plants when they were searched for thoroughly. Nearly 60 Indian endemic species have been rediscovered in this way during 1990–98, for example *Cynometra bourdillonii*, *Dialium travancoricum*, *Humboldtia bourdillonii*, *Inga cynametroides*, *Taeniophyllum scaberulum* and *Aenhenrya rotundifolia*. Some papers report the rediscovery of more than a dozen 'extinct' species.

These rediscoveries are directly related to the amount of collecting effort invested, and we believe that most tropical (rather than temperate) 'extinction' is actually non-availability of data rather than genuine extinction. No Indian biologist has yet questioned the credibility of these reports. We suggest that the Indian plants classified as extinct because they have not been collected recently should instead be termed 'plants to be rediscovered'. This may save the word 'extinction' from constant misuse.

Pointing out examples of plants that have been considered extinct but then re-established from seed banks, Moore suggested⁵ that monitoring of the soil seed or spore bank should be mandatory before

designating a plant as extinct. But is this possible in a country such as India, where the search for the mature plant itself is a difficult task; where biodiversity documentation is incomplete; and where taxonomists themselves are 'critically endangered' owing to lack of funding⁶?

K. P. Rajesh, P. V. Madhusoodanan

Department of Botany, Calicut University, Kerala 673 635, India

1. Moore, P. D. *Nature* **392**, 661–662 (1998).
2. Rumsey, F. J., Jermy, A. C. & Sheffield, E. *Watsonia* **22**, 1–19 (1998).
3. Gadgil, M. & Meher-Homji, V. M. in *Conservation in Developing Countries: Problems and Prospects* (eds Dannel, J. C. & Serrao, J. S.) 175–197 (Bombay Natural History Society, 1990).
4. Jain, S. K. & Sastry, A. R. K. in *Biological Aspects of Rare Plant Conservation* (ed. Synge, H.) 59–66 (Wiley, New York, 1981).
5. Moore, P. D. *Nature* **303**, 572 (1983).
6. Khoshoo, T. N. *Curr. Sci.* **69**, 14–17 (1995).

Proposed GMO rules lack scientific sense

Sir — Your article about negotiations on the global regulation of genetically modified organisms (GMOs) omitted much of the essential context of this ill-conceived undertaking (*Nature* **398**, 6; 1999).

It is misleading to represent the lack of agreement at the talks as a "conflict between trade and environmental concerns". 'Pseudo-environmental concerns' would be more apt. The scientifically insupportable scope of the proposed biosafety protocol focuses regulatory attention on experiments of largely negligible risk. It

covers GMOs, but no other organisms, no matter how pathogenic or otherwise dangerous to the environment.

Ironically, the protocol would burden with unscientific, expensive and unnecessary regulation environment-friendly products that can be produced by recombinant DNA technology, and which are needed by developing countries.

US negotiators at the talks in Cartagena could have argued persuasively that the proposed regulations lack scientific and common sense, but their position instead focused exclusively on trade considerations, aiming to protect agribusiness interests.

The United Nations' proposed protocol would make GMOs artificially expensive to test, produce and use. According to a US Department of Agriculture study, the prices of wheat and coarse grains (corn, barley and sorghum) could increase worldwide by an average of 2 per cent and 5.6 per cent, respectively. Developing countries would spend more on food and be prevented from participating in technological trends.

Future talks should be based on scientific principles, actual product risk and the public interest, rather than politics, expediency and narrow self-interest.

Henry I. Miller

Hoover Institution,
Stanford, California 94305-6010, USA

Others should follow the US line on bioweapons

Sir — Your article on the controversy surrounding destruction of the smallpox virus presented well the difficult and complicated issues facing international policy-makers (*Nature* 398, 741; 1999). But one statement attributed to me was seriously in error. It is alleged that I said that the US Department of Defense wished to see stocks of variola virus retained in order to retaliate in kind should the United States be subjected to a bioweapons attack. From this, it might be implied that the United States was not intent on meeting the obligations of the Biological Weapons Convention that ban offensive weapons. That would be a serious matter indeed.

The US stock of offensive weapons was destroyed more than 20 years ago. During my years of government service, I know of no one who ever suggested using biological weapons in retaliation or in any other manner. One would hope that this might one day be the accepted norm for all countries.

D. A. Henderson

Center for Civilian Biodefense Studies, Johns Hopkins University, Candler Building, Suite 850, 111 Market Place, Baltimore, Maryland 21202-6907, USA

There's still a place for physics out west

Sir — Tony Reichhardt's article, "Wyoming physics faculty faces closure", grossly overstates our recommendations on the future of the Department of Physics and Astronomy at the University of Wyoming (*Nature* 398, 357; 1999). Since the article appeared, we have further refined our recommendations, which will clarify the future of physics instruction.

It was never suggested that the university should not teach physics. The academic plan did, however, raise serious questions about the health of the Department of Physics and Astronomy. These in turn raise questions about the number of degree programmes we will try to support. Currently, we offer bachelors, masters and doctoral degrees in physics, and a bachelors degree in astronomy/astrophysics.

Physics and astronomy is a troubled department. Last year, with a budget of more than \$1 million, it produced only one graduate with a baccalaureate degree. We have an obligation to ask whether a budget of this magnitude is justified.

There are additional problems. Our recommendation to eliminate the department's graduate programmes and to question the future of the baccalaureate degree was based largely on a 1998 peer review of the mathematics and physical sciences divisions within the university's College of Arts & Sciences.

The reviewers, representatives from other University of Wyoming colleges as well as from other regional universities, examined six departments. Five were rated 'very good' or 'acceptable'; only physics and astronomy was rated 'unacceptable'.

The review found that, between 1992 and 1996, undergraduate enrolment fell by 44 per cent, graduate enrolment by 29 per cent, and degrees granted by 53 per cent. Enrolment in many undergraduate courses was lower than the university's standard minimum. About half the faculty members were inactive in research. Collegiality among at least some faculty members was low, and communication was a major problem. The Wyoming Infrared Observatory (WIRO) appeared to be underused and underequipped, owing in part to the staff's inability to communicate and work together.

Both this review and an evaluation of WIRO by leading astronomers also indicated that changes must be made in the administration of the observatory, to improve its cost-effectiveness, and in the quality of its instruments to improve the science. Given the many calls on this

institution's limited funds, it was imperative that the academic plan should question whether this is a sensible continuing investment. To treat any discipline as a sacred cow would make a mockery of any planning process.

Since the article in *Nature*, we have completed our second draft of the plan. We received thousands of comments on the draft. Many comments dealt with the Department of Physics and Astronomy.

The comments did not change our view that the department is troubled. However, we were convinced that we could not have a credible department without offering at least a bachelors degree in the subject.

The new draft plan says the university's intention is to maintain a baccalaureate programme in physics, but it also places the responsibility where it belongs — on the faculty members — to rebuild the programme, beginning with the undergraduate programme. Graduate students will be able to complete their studies, but the plan suggests a moratorium on accepting new graduate students. Whether the masters or doctoral programmes are restored in the future will depend upon the progress made at undergraduate level. We will continue to look at new creative ways to manage the underused infrared telescope.

We believe that the people of Wyoming want a stronger, more focused university. We asked tough questions about every programme, including physics. The article raised the spectre of whether Wyoming would be shamed by being the only state-run university without a physics degree programme. In fact, we will maintain our baccalaureate programme and will rebuild the department, but only progressively and in response to specific benchmarks of performance. We know of no other way to preserve the quality educational experience we offer our students.

Philip L. Dubois

(President)

Tom Buchanan

(Vice-President for Academic Affairs)

University of Wyoming,
Laramie, Wyoming 82071, USA

Reichhardt replies — My News report said, in fact, that the university would still teach physics, and that not offering the bachelor's degree was only an option. I wrote: "The plan recommends that consideration be given to scrapping the bachelors degree programme and offering physics courses only as a 'service function'..." The university later chose to keep the bachelor's degree, but it was still a possibility that it would be killed off when the article was written.

Tony Reichhardt

104 Cleremont Drive, Fredericksburg,
Virginia 22405, USA

Scientific societies build better nations

Scientific societies in the developing world must take a stake in their countries' future. They should be proactive in fostering a culture supportive of economic development driven by science and technology.

Leo Tan Wee Hin and R. Subramaniam

In many developing countries, science is treated as a marginal activity, or even as an adornment, to quote the Pakistani Nobel prizewinning physicist Abdus Salam. It is, then, hardly surprising that the tradition of scientific societies has not taken deep root. Very few are as active as their Western counterparts, and fewer still have international stature. This partly explains why many developing countries have been slow to gain economic benefit from advances in science and technology.

Scientific societies are deeply embedded in Western culture. In the two centuries following the establishment, in 1660, of the pioneering Royal Society in the United Kingdom, most of the countries of Europe and North America followed suit. These non-governmental organizations — professional bodies with altruistic objectives — have worked selflessly to promote the public understanding of science and raise the status of scientific disciplines and professions. The evolution of a scientific culture in the West has been due in no small measure to the efforts of these societies. They are highly regarded by the establishment and their views command respect. The societies are also vigorous publishers of scientific journals, contributing to the quality of research.

In the developing world, the desired levels of economic success are still elusive, despite prudent initiatives such as encouraging foreign investment, promoting industrialization, instituting macroeconomic reforms, and emphasizing research and technological education. With the approach of the third millennium, and the recognition that science and technology are agents for change, we must ensure that developing countries are not further isolated from the global economic mainstream. The scientific community has a responsibility in this regard.

The Western experience shows that scientific societies in developing countries have much to contribute to nation building. In the case of Singapore, the scientific community recognized the importance of scientific societies in nation building soon after independence in 1965. Since then, the Singapore National Academy of Science and its scientific societies have been proactive in complementing the government's efforts in the cause of development. They promote numerous annual programmes aimed at a national audience, including science festivals, science olympiads, mathematics olympiads, science and technology competitions, badge



Societies promote festivals and science and maths olympiads... spreading a culture supportive of science among the public

schemes for student project work in 12 disciplines, lectures and many other activities.

The effectiveness of these programmes in spreading a culture supportive of science and technology among the public in general and students in particular is well acknowledged. Funding has been a major but not insurmountable issue in implementing awareness programmes. The problem of scarce funding has been overcome by pooling expertise and resources and cultivating a spirit of commitment, enthusiasm and volunteering among society members. Alliances with government and private corporations have been forged. In parallel with its public programmes, the scholarly activities of the academy and its societies have raised the profile of the disciplines and professions they espouse.

It seems to us that scientific societies in developing countries are well positioned to help bring about a culture supportive of science- and technology-driven development through public initiatives and other measures. Representing a concentration of expertise and intellectual resources, they are in a position to take an equity stake in their countries' futures. Although a lack of funds is one reason why such societies have not

maximized their potential, this can be more than made up for by partnerships with government organizations and private corporations, as well as by tapping the large workforces of these countries. This would have the effect of strengthening existing institutions and enhancing their stature.

The success of general scientific societies in the West has led to the proliferation of more specialized societies catering to diverse interests and objectives. Many have regional branches, and these have helped to enhance their reach and effectiveness. The scientific intelligentsia in developing countries can emulate such measures.

More importantly, scientific societies can be established without the need for institutional infrastructure, official sanction or massive funding. It just requires the coming together of a core group of scientists committed to translating a vision into reality. Nearly three decades after its birth, the Singapore National Academy of Science and its affiliates have still not found it necessary to have permanent premises, despite their wide range of activities.

There needs to be more support for scientific societies from international aid agencies such as the United Nations Educational, Scientific and Cultural Organization, the World Bank, the International Monetary Fund and the rich countries that provide aid to the Third World. Indeed, such agencies should give modest annual grants to learned societies in developing countries. Over 70% of the support for science and technology in Africa comes from foreign aid. Even if only a small percentage of this went to scientific societies, it would be a tremendous help. Support and encouragement would establish new partnerships, thereby catalysing development. It would also further energize the scientific intelligentsia in developing countries for a noble cause, and accelerate the creation of institutions of creditable standing out of such scientific societies.

As Salam noted, "It is basically mastery and utilization of science and technology that distinguishes the South from the North." It is time for scientific societies in developing countries to help to close that gap by taking on a higher profile. □

Leo Tan Wee Hin, president of the Singapore National Academy of Science, is at Nanyang Technological University, 469 Bukit Timah Road, Singapore 259756.

R. Subramaniam, honorary secretary of the Singapore National Academy of Science, is at Singapore Science Centre, 15 Science Centre Road, Singapore 609081.

DAVID NEWTON

Cultural primatology comes of age

Frans B. M. de Waal

The chimpanzee keeps inching closer to humanity. After decades of patiently gathering information, the heads of seven field-sites pool their knowledge to reveal the astonishing variation in tool technology and social customs in chimpanzees across Africa.

The question of whether animals have culture is a bit like asking whether chickens can fly. Compared with an albatross or a falcon, perhaps not, but chickens do have wings, they do flap them, and they can get up in the trees. Similarly, viewed from the cultural heights achieved by humans in art, cuisine, science and politics, other animals seem to be nowhere in sight. But what if we change perspective, and don't measure them by our standards? This is what Kinji Imanishi, a Japanese anthropologist, proposed in the early 1950s¹. In an imaginary debate between an evolutionist, a layman, a monkey and a wasp, Imanishi suggested that culture — defined as the non-genetic transmission of habits — was entirely possible, and even likely, for animals other than humans.

This insight prepared the minds of primatologists for the spread of potato washing among macaques on the tiny Japanese island of Koshima (Fig. 1). One juvenile female pioneered the habit of carrying sweet potatoes to the water to clean off the dirt. Her mother and closest peers soon followed, and the habit spread to others. Within a decade, the whole of the population under middle age was washing potatoes². But Western anthropologists and psychologists were distinctly uncomfortable with application of the 'culture' label to mere monkeys. This led to definitions of culture that required linguistic mediation, a high speed of behavioural change, or full-blown imitation^{3–5}. Soon the debate shifted to the question of whether monkeys and apes actually 'ape'. In drawing such lines, the critics echoed earlier views that designated our species as the only one to possess culture⁶.

If animal groups vary in a single behaviour such as potato washing, there is, perhaps, not much reason to use a loaded term such as 'culture' — 'group-specific trait' or 'tradition' should do instead. But the first intimation that things might not be so simple with regard to our closest relatives came in 1992, with William McGrew's review of chimpanzee tool-use in the wild⁷. Since then,

new observations have appeared, one by one, in the primatological journals. And now, on page 682 of this issue, comes the grand synthesis by Andrew Whiten and his colleagues⁸, describing the various habits of chimpanzees at no less than seven well-established field sites. The record is so impressive that it will be hard to keep these apes out of the cultural domain without once again moving the goalposts.

The evidence comes from a survey of all suspected cultural variants in wild chimpanzees, including behaviour patterns never published before. Some populations, for example, fish for ants with short sticks, eat-

ing the prey off the stick one by one. But at least one population has developed the more efficient technique of accumulating many ants on a long wand, after which all insects are swept into the mouth with a single hand motion.

After compiling a first list, Whiten *et al.*⁸ rated behaviour patterns on a scale from customary to absent at each field site, and the ecology of each site was taken into account. For instance, chimpanzees will not sleep in ground nests (as opposed to tree nests) at sites with high leopard or lion predation. Such ecologically explainable differences were excluded from the list, leaving no fewer than 39 behaviour patterns — far more than reported for any other animal — that vary across chimpanzee communities in Africa.

Genes determine general abilities, such as tool use, but it is hard to imagine that they instruct apes how exactly to fish for ants or whether or not to make cushy seats out of vegetation. Moreover, Whiten and colleagues found no evidence that habits vary more between, than within, the three existing subspecies of chimpanzee. So genetics cannot account for the observed variability. Take the two best-known communities: those studied by Jane Goodall in Gombe National Park, and by Toshisada Nishida in the Mahale Mountains, both in Tanzania. The two sites are only 170 km apart, and both are inhabited by the Eastern subspecies



Figure 1 Japanese monkey washing potatoes. About half a century ago, a juvenile Japanese macaque developed sweet-potato washing on the island of Koshima. The habit spread to the rest of the population. None of these monkeys is still alive today, but their descendants are still washing potatoes.



Figure 2 Transmitting culture — this chimpanzee shares the nut that she has just cracked open. Nut-cracking, one of the best-studied cultural variants, is acquired by young chimpanzees only after many years of practice.

F. B. M. DE WAAL

C. BOESCH

(*Pan troglodytes schweinfurthii*). In Mahale, grooming apes customarily clasp two hands together high above their heads — something never seen in 40 years at Gombe⁹.

Studies of chimpanzees in captivity support the emerging picture of cultural apes. Because captive groups are relatively young, new habits often develop and their spread can be carefully charted¹⁰. Also, new techniques can be demonstrated to the apes by human experimenters, to see how faithfully they are copied¹¹. All in all, the evidence is overwhelming that chimpanzees have a remarkable ability to invent new customs and technologies, and that they pass these on socially rather than genetically (Fig. 2).

The definition of culture will no doubt keep changing, but Whiten *et al.*⁸ rightly take the position, common in the life sciences, that mechanisms are of secondary importance. In the same way that the definition of respiration doesn't specify whether the process takes place through skin, lungs or gills, the concept of cultural propagation does not specify whether it rests on imitation, teaching or language. The 'culture' label

befits any species, such as the chimpanzee, in which one community can readily be distinguished from another by its unique suite of behavioural characteristics. Biologically speaking, humans have never been alone — now the same can be said of culture. □

Frans B. M. de Waal is at *Living Links, The Yerkes Regional Primate Research Center and Department of Psychology, Emory University, Atlanta, Georgia 30329, USA.*
e-mail: dewaal@emory.edu

1. Itani, J. & Nishimura, A. in *Precultural Primate Behavior* (ed. Menzel, E. W. Jr) 26–50 (Karger, Basel, 1973).
2. Watanabe, K. in *The Ethological Roots of Culture* (eds Gardner, R. A., Chiarelli, A. B., Gardner, B. T. & Plooj, F. X.) 81–94 (Kluwer, Dordrecht, 1994).
3. Galef, B. G. *Hum. Nature* **3**, 157–178 (1990).
4. Premack, D. & Premack, A. J. in *Companion Encyclopedia of Anthropology* (ed. Ingold, T.) 350–365 (Routledge, London, 1994).
5. Tomasello, M., Kruger, A. C. & Ratner, H. H. *Behav. Brain Sci.* **16**, 495–552 (1993).
6. Kroeber, A. L. & Kluckhohn, C. *Pap. Peabody Mus. Am. Archeol. Ethnol.* **47**, 41–72 (1952).
7. McGrew, W. C. *Chimpanzee Material Culture: Implications for Human Evolution* (Cambridge Univ. Press, 1992).
8. Whiten, A. *et al. Nature* **399**, 682–685 (1999).
9. McGrew, W. C. & Tutin, C. E. G. *Man* **13**, 243–251 (1978).
10. de Waal, F. B. M. & Seres, M. *Am. J. Primatol.* **43**, 339–346 (1997).
11. Whiten, A. *J. Comp. Psychol.* **112**, 270–281 (1998).

Enzymes

Picking a winner

Roger Sheldon

Enzymes are remarkable catalysts that perform their tasks with alacrity in water at room temperature, and with a high degree of regioselectivity (for example, substitution at a particular position in an aromatic ring) and enantioselectivity (for a particular optical isomer). Because of the trend towards more environmentally friendly processes, and drugs that are more effective and cause fewer side effects (such as optically pure isomers), biocatalysis is rapidly becoming an attractive option for the manufacture of fine chemicals. The industrial application of certain enzymes, for example most hydrolytic enzymes, is relatively straightforward, whereas for others, such as the reduction–oxidation, or redox, enzymes, the need for enzyme cofactors makes their use more problematic. Nonetheless, it is precisely these enzymes that mediate the most interesting synthetic conversions.

A major challenge in biocatalysis is to harness the enormous potential of the ubiquitous cytochrome P450 monooxygenases¹. These enzymes play key roles in the biosynthesis of prostaglandins and steroids, among others, and in the detoxification of foreign substances in the body, including drugs, pesticides and petroleum hydrocarbons. They are promiscuous catalysts, using molecular oxygen to insert an oxygen atom into a number of substrates. Many of these conversions are of potential industrial interest, for example in the production of pure enantiomers

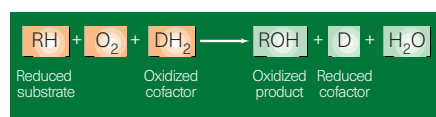


Figure 1 Oxidation–reduction reaction catalysed by cytochrome P450 monooxygenases.

of epoxides. Unfortunately, they suffer from several drawbacks — low stability and turnover rates and the need for a complex cofactor regeneration system — that prohibit their industrial application. The paper by Arnold and co-workers on page 670 of this issue² constitutes an important step towards the ultimate goal of commercial viability. They successfully apply the technique of 'directed evolution', using random mutagenesis and DNA shuffling, to develop mutants of a bacterial cytochrome P450 enzyme that efficiently use hydrogen peroxide (H₂O₂) as the source of oxygen, thereby bypassing the need for cofactor regeneration.

Cytochrome P450 monooxygenases catalyse redox reactions, which involve taking one oxygen atom from O₂ and inserting it into a substrate (hence the name monooxygenase). The second oxygen atom is reduced to water (Fig. 1). Cytochromes have an iron-containing porphyrin group (a haem complex) that catalyses electron-transfer processes by cycling between oxidized and reduced forms of iron. Cytochrome P450 (the name is derived from the fact that its spectrum exhibits a characteristic peak

around 450 nm) differs from other cytochromes in that it uses O₂ and a reducing source to oxidize substrates. The necessary reducing source is provided by the cofactor system, which cycles between D and DH₂ (Fig. 1).

The cofactor, a nicotinamide adenine dinucleotide (NAD or NADPH), is regenerated *in vivo* by a complex system involving multiple redox proteins, the reducing agents being ultimately derived from an energy source such as glucose. The need for a complex cofactor regeneration system, along with the poor stability of these enzymes outside the cell, dictates that fermentation is the best way to perform such reactions in practice. But the complexity of fermentation processes limits the application of these enzymes to expensive products such as steroids. Their scope could be expanded if it were possible to use the isolated enzymes. Indeed, the need for cofactor regeneration can be circumvented altogether by using H₂O₂ as the oxygen source in the so-called 'peroxide shunt' pathway¹ (Fig. 2). Unfortunately, this is not very efficient, and in order to be synthetically viable would have to be improved considerably. This was the starting point for the study by Arnold and co-workers².

After aeons of evolution, through mutation and Darwinian selection, enzymes are ideally suited to the tasks they perform *in vivo*. They have not evolved for the purpose of producing chemicals on an industrial scale, and so often lack the necessary features, such as stability outside the cell and high turnover rates. But, using the tools of modern biotechnology, it is now possible to mimic the evolution of enzymes *in vitro* — so-called directed evolution — in weeks rather than millions of years. The advent of error-prone polymerase chain reaction (PCR) has made it possible to selectively target the gene encoding a particular enzyme and generate a library of genes containing point mutations. Recombination produces further mutations that can be screened to find variants leading to superior enzymes³. It is a relatively easy task to make thousands of mutants, but the real key to success is developing an assay that can pick out the useful ones.

For their directed evolution study Arnold and co-workers² chose the widely studied cytochrome P450_{cam} from the bacterium *Pseudomonas putida*, the structure of which has been determined by X-ray diffraction. This enzyme catalyses the hydroxylation of camphor *in vivo*, showing only weak activity towards naphthalene — a component of coal tar used to manufacture dyes and resins. The hydroxylation of naphthalene with H₂O₂ was chosen as the model reaction to be studied, presumably because a bicyclic aromatic leads to more highly conjugated, and therefore more highly coloured, products.

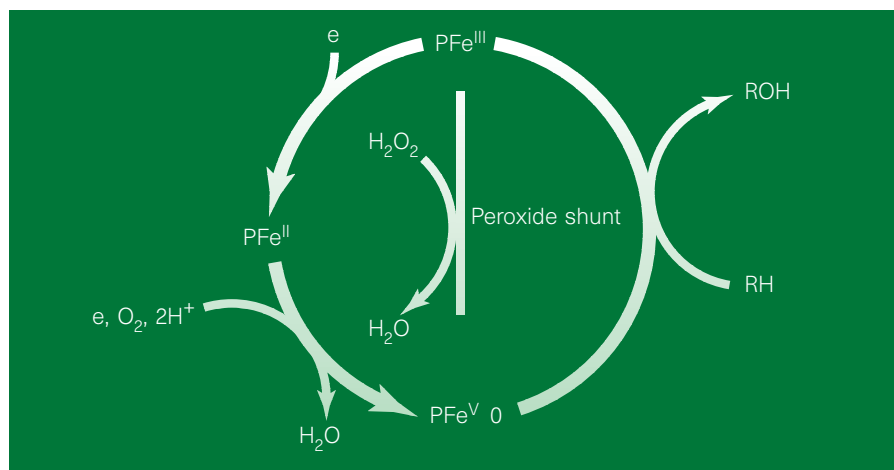


Figure 2 Catalytic cycle of cytochrome P450. Arnold and co-workers² have developed more efficient enzymes using the 'peroxide shunt' pathway to bypass a complex cofactor regeneration system. In this pathway, reaction of hydrogen peroxide (H_2O_2) with the oxidized iron porphyrin complex (PFe^{III}) leads to the active oxidant ($\text{PFe}^{\text{V}}=\text{O}$) in a single step, instead of by reduction to PFe^{II} followed by reaction with O_2 and reducing sources (electrons and protons) from the cofactor. Reaction of the substrate (RH) with $\text{PFe}^{\text{V}}=\text{O}$ gives the product (ROH), returning the enzyme to the resting state (PFe^{III}).

The main task was to develop a suitable high-throughput screening method. The authors achieved this in the following novel way. The enzyme horseradish peroxidase catalyses the oxidative coupling of hydroxylated aromatics, using H_2O_2 as the oxidant, to generate coloured and fluorescent materials. They showed that expression of cytochrome P450_{cam} with a variant of horseradish peroxidase (HRP1A6) in *Escherichia coli* creates a pathway for the conversion of aromatic substrates such as naphthalene into fluorescent compounds.

Approximately 200,000 mutants of cytochrome P450_{cam} , produced by error-prone PCR and coexpressed with HRP1A6, were screened for activity in the reaction of naphthalene with H_2O_2 using fluorescence digital imaging. A benefit of this screening method is that different hydroxylation substitution patterns lead to different colours. For instance, a combination of 1- or 2-naphthol and 2,7-dihydroxynaphthalene produces red fluorescence, whereas 2-naphthol with 1,5-dihydroxynaphthalene produces yellow fluorescence. So, the screening allows insight into the altered regiospecificities of the various mutants. Indeed, a palette of colours was observed with bacteria expressing the P450_{cam} mutants (see Fig. 4 on page 672).

Three brightly fluorescing clones were selected for growth and their enhanced activity was confirmed in a whole-cell assay. Moreover, DNA sequencing revealed that all three of the improved P450 mutants share the same mutation (a lysine substituted for glutamate). There is no obvious, rational explanation for this mutation, which highlights the unique ability of the directed evolution technique to identify mutants that may not be selected by other methods. A second-generation P450_{cam} library prepared by

in vitro recombination of five improved variants yielded two clones that were roughly 20 times as active as the wild-type P450_{cam} in the whole-cell assay. In short, Arnold and co-workers² have developed an elegant screening method that enables the directed evolution of cytochrome P450 enzymes to generate mutants that are more active, possibly

more stable and bypass the need for cofactor regeneration in the natural enzyme.

The resulting enzymes are still far from being commercially viable, yet the first hurdle has been passed. The enzymes may show improved activities in other characteristic P450 transformations, such as olefin epoxidation. But there may be a limit to the stability that can be achieved with haem-dependent enzymes, owing to the inherent instability of the porphyrin structure under oxidizing conditions (a lesson we have learned from studies of peroxidases⁴). So it would be interesting to extend the methodology of Arnold and co-workers to the directed evolution of non-haem monooxygenases, some of which catalyse similar reactions to the P450 enzymes. Perhaps this will eventually lead to commercially viable monooxygenases for the regioselective and enantioselective oxidation of fine chemicals. □

Roger Sheldon is in the Department of Organic Chemistry and Catalysis, Delft University of Technology, Julianalaan 136, 2628 BL Delft, The Netherlands.

e-mail: r.a.sheldon@stm.tudelft.nl

- Sheldon, R. A. (ed.) *Metalloporphyrins in Catalytic Oxidations* (Dekker, New York, 1994).
- Joo, H., Lin, Z. & Arnold, F. H. *Nature* **399**, 670–673 (1999).
- Stemmer, W. P. *Nature* **370**, 389–391 (1994).
- van Deurzen, M. P. J., van Rantwijk, F. & Sheldon, R. A. *Tetrahedron* **53**, 13183–13220 (1997).

Photonics

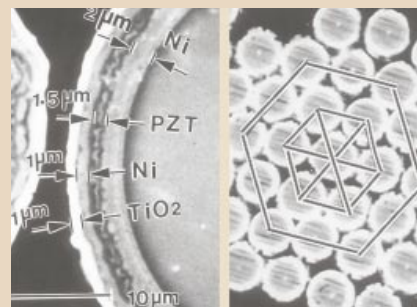
Crystals to order

Controlling the behaviour of particles suspended in a fluid is critical, whether you are trying to mix paint or separate crude oil. A group at the Hong Kong University of Science and Technology are taking the control of particle properties in solution to a new level, in the hope of creating useful photonic materials (the optical analogues of semiconductors).

Ping Sheng and his colleagues (*Phys. Rev. Lett.* **82**, 4248–4251; 1999) coated micrometre-sized glass spheres with a layer of nickel, followed by PZT (lead zirconate titanate, a quartz-like ceramic), another layer of nickel and finally titanium oxide. The four coatings (left side of figure) are intended to give a large response to applied electric or magnetic fields.

The coated spheres are suspended in silicon oil. In zero applied field, the spheres are randomly dispersed. As the external electric field is increased, the particles form columns, and other changes emerge when a magnetic field is applied. By freezing the system at different field strengths and taking cross-sectional micrographs, Sheng and co-workers have snapshots of the changing structure (right side of figure).

They discover that as the magnetic field



increases, the spheres change from a body-centred-tetragonal (bct) structure to a face-centred-cubic (fcc) structure. Small movements of the spheres produce the structural transition, without any long-range diffusion. Such a transition can be achieved — simply by varying the relative strengths of the external fields — because the difference in energy between the bct and fcc structures is very small.

A tunable crystal structure might offer new ways to make photonic materials that permit the passage of photons of certain energies, while excluding others. The difficulty in finding photonic crystals with the right properties (lattice dimensions from a few hundred nanometres to tens of micrometres) could be solved by these made-to-order crystals.

Sarah Tomlin

DNA repair

Variants on a theme

Richard D. Wood

Suppose you are travelling on a long car journey and come to a deep pothole in the middle of the road. You can either terminate your travels, wait for the hole to be repaired, or plunge straight across it. The last course of action is risky, but it might at least allow you to finish the trip. A similar decision has to be made when, as cells are duplicating their DNA, the replication apparatus encounters a damaged site that blocks further progress. But a dramatic development, reported by Hanaoka and colleagues¹ on page 700 of this issue, reveals a mechanism that human cells use to bypass DNA damage instead of removing it. Remarkably, people who lack the enzyme that carries out one type of bypass have an inherited, cancer-prone syndrome.

Damage to a cell's genome constantly arises by the natural decay of DNA (in aqueous solution or by oxygen free-radical damage²) or by damage from external agents such as γ -rays and UV light. The result is lesions, usually in one strand of the double helix, which block progression of the enzymes that replicate DNA (the DNA polymerases). We often think of cell-cycle checkpoints — stages at which the cell-division cycle can

stall — as a mechanism for cells to repair any damage in their DNA before DNA replication or division is complete. But this is not the whole story. DNA damage does not always need to be removed, and cells can divide with tremendous amounts of damage in their genomes³. How can they achieve such damage tolerance?

Hanaoka and co-workers¹ have solved an important piece of this problem by examining a rare, inherited, cancer-prone syndrome called the xeroderma pigmentosum 'variant', or XP-V. Besides the XP-V form, the disorder xeroderma pigmentosum has seven other 'classical' genetic complementation groups, designated XP-A through to XP-G. The seven genes responsible encode subunits of the DNA nucleotide excision repair machinery, which removes UV-induced lesions and other damage from the genome⁴. The XP-V form differs in that sufferers have normal nucleotide excision repair, and their cells are only slightly more sensitive than normal to UV light. Nevertheless, like other patients with xeroderma pigmentosum, they have a very high risk of sunlight-induced skin cancer⁵.

An early hint that XP-V cells have diffi-

culty in bypassing — rather than removing — DNA damage came from work by Lehmann and colleagues⁶. This team showed that when XP-V cells are irradiated with UV light, they are exceptionally slow to join together new DNA fragments during DNA replication. The XP-V cells were subsequently found to accumulate a much higher frequency of UV-induced mutations in their genome than normal. Moreover, the main mutational changes were different from those that occur in normal cells⁷.

To examine the XP-V defect more closely, Hanaoka and colleagues first measured the ability of DNA containing a single UV-induced lesion (a thymine dimer) to be replicated in soluble extracts from human cells. The XP-V cell extracts were very poor at replicating past the lesion, confirming the findings of other groups^{8–11}. The authors then purified an activity from normal cells that allowed XP-V cell extracts to bypass DNA lesions. They identified this activity — which is described in the latest *EMBO Journal*¹² — as a DNA polymerase that works with remarkable accuracy. During replication, it inserts only adenine residues, the complementary DNA base, opposite the thymine residues of a UV-induced thymine dimer.

The latest step¹ has been to sequence the purified DNA polymerase and clone the corresponding gene. Hanaoka and colleagues now reveal that it is, in fact, the human version of DNA polymerase η — a translesion-bypass polymerase that has already been identified in budding yeast¹³. The authors found that purified recombinant human DNA polymerase η could correct the bypass defect of XP-V cell extracts. Moreover, all XP-V cells examined had mutations in the DNA polymerase η gene. So, in one swoop, the biochemical defect in the last unsolved group of xeroderma pigmentosum has been exposed. It is also the first example of a human DNA polymerase that bypasses UV-induced lesions efficiently in a relatively error-free manner.

The protein sequence of DNA polymerase η places it firmly into an emerging family of bypass DNA polymerases. These include the yeast equivalent¹³, as well as UmuC (a component of the UmuD₂C complex, which facilitates bypass of DNA damage in *Escherichia coli*). The new discovery also means that there are now at least seven DNA polymerases in mammalian cells: DNA polymerases α , β , δ and ϵ are involved in DNA replication and repair; DNA polymerase γ is found in mitochondria; and DNA polymerase ζ increases the frequency of UV-induced mutagenesis¹⁴.

DNA polymerase η is entirely eliminated in many patients with XP-V, indicating that this protein is not essential for life. In normal cells, many UV-induced thymine dimers are bypassed in an error-free fashion by DNA polymerase η (Fig. 1). Others are bypassed in

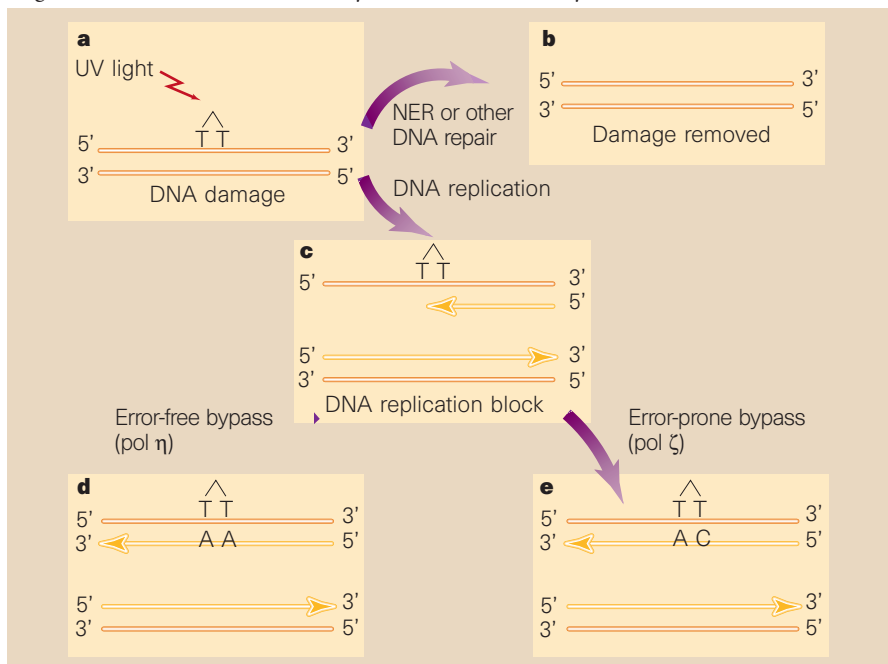


Figure 1 Model for DNA synthesis across sites of damage. a, Ultraviolet (UV) light and other mutagens create damage in DNA, such as the thymine dimer depicted. b, Often such damage is removed, by nucleotide excision repair (NER), for example. The NER process is defective in most xeroderma pigmentosum (XP) groups, but not in the XP variant (XP-V). c, If not repaired, the thymine dimer blocks progression of the DNA replication fork. d, Bypass of the dimer facilitated by DNA polymerase η (pol η) is usually free of errors if the correct residues are inserted opposite the lesion. Hanaoka and colleagues¹ have shown that this process is defective in XP-V cells. e, Bypass of the dimer facilitated by DNA polymerase ζ (pol ζ) is error prone, as incorrect bases are often inserted. This can still occur in XP-V cells.

an error-prone fashion, by DNA polymerase ζ , for example, often leading to mutations. XP-V cells are much more sensitive than normal to mutation. This can be thought of as a 'channelling' of UV-induced lesions into more error-prone pathways, such as that controlled by DNA polymerase ζ .

But why do cells have an error-prone damage-bypass pathway at all, given that they already have the almost error-free DNA polymerase η pathway? Perhaps error-prone bypass is reserved for lesions that are particularly difficult to cross — because of either their chemical structure, or their particular environment in the genome. So far, however, bypass has been studied for a limited number of lesions, and only the thymine dimer has been looked at in detail.

Clinically, xeroderma pigmentosum is very similar in the XP-A to XP-G (nucleotide excision repair-deficient) and XP-V (DNA polymerase η -deficient) forms. What does this mean? In XP-A to XP-G cells, many lesions remain unrepaired. Most lesions result in cell death, but some are bypassed by translesion DNA synthesis, using either error-prone or error-free mechanisms. Lesions bypassed by the error-prone pathway often result in mutations, explaining the high incidence of skin cancer. Although XP-V cells have fewer UV-induced lesions, because nucleotide excision repair can operate, unrepaired lesions are more likely to be

bypassed by error-prone mechanisms owing to inactivation of the error-free pathway. Moreover, thymine dimers with 'mismatches' opposite them are repaired more efficiently by nucleotide excision repair than are correctly paired dimers¹⁵. So the intact nucleotide excision repair in XP-V might even accelerate mutagenesis. In the end, though, the effect of the accumulated mutations is the same — it increases the chance of generating cells that are set on the road to tumour formation. □

Richard D. Wood is at the Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche Lane, South Mimms, Hertfordshire EN6 3LD, UK.

1. Masutani, C. *et al.* *Nature* **399**, 700–704 (1999).
2. Lindahl, T. *Nature* **362**, 709–715 (1993).
3. Spivak, G. & Hanawalt, P. C. *Biochemistry* **31**, 6794–6800 (1992).
4. de Laat, W. L., Jaspers, N. G. J. & Hoeijmakers, J. H. J. *Genes Dev.* **13**, 768–785 (1999).
5. Cleaver, J. E. & Kraemer, K. H. in *The Metabolic Basis of Inherited Disease* 7th edn (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 4393–4419 (McGraw-Hill, New York, 1995).
6. Lehmann, A. *et al.* *Proc. Natl Acad. Sci. USA* **72**, 219–223 (1975).
7. Wang, Y.-C., Maher, V. M., Mitchell, D. L. & McCormick, J. J. *Mol. Cell. Biol.* **13**, 4276–4283 (1993).
8. Cordeiro-Stone, M. *et al.* *Biol. Chem.* **272**, 13945–13954 (1997).
9. Svoboda, D. L., Briley, L. P. & Vos, J. M. H. *Cancer Res.* **58**, 2445–2448 (1998).
10. McGregor, W. G., Wei, D., Maher, V. M. & McCormick, J. J. *Mol. Cell. Biol.* **19**, 147–154 (1999).
11. Cordonnier, A. M., Lehmann, A. R. & Fuchs, R. P. P. *Mol. Cell. Biol.* **19**, 2206–2211 (1999).
12. Masutani, C. *et al.* *EMBO J.* **18**, 3491–3510 (1999).
13. Johnson, R. E., Prakash, S. & Prakash, L. *Science* **283**, 1001–1004 (1999).
14. Gibbs, P. E. *et al.* *Proc. Natl Acad. Sci. USA* **95**, 6876–6880 (1998).
15. Mu, D. *et al.* *Mol. Cell. Biol.* **17**, 760–769 (1997).

Comets

Putting the CO in coma

Jacques Crovisier

When comet Hale-Bopp was at its closest to the Sun (roughly 0.9 AU, where 1 AU = 1 astronomical unit, the Earth-to-Sun distance) at the beginning of April 1997, it released about 300 tons of water and 100 tons of carbon monoxide (CO) per second. Whereas cometary water is known to arise from the sublimation of ices in the nucleus, the origin of cometary CO is more controversial. On page 662 of this issue, DiSanti *et al.*¹ report infrared observations of CO emission from comet Hale-Bopp over a range of distances from the Sun. They reveal that cometary CO has a dual origin — 50% from the nucleus and 50% from a distributed source in the cometary atmosphere. Moreover, they show that the distributed source switches off at a certain distance from the Sun, suggesting a thermal production process and giving some clues to its identity.

Within the 'dirty snowball' model, first proposed by Fred Whipple in 1950, comet nuclei are thought to consist of ices — mainly frozen water, with some carbon dioxide, methane and ammonia — mixed up with dust particles. When these icy cores

approach the Sun, radiation causes the ices to vaporize, releasing a diffuse envelope of dust and gas called the coma, and several tails. The comet nuclei are too small to be observed directly, but the coma can be millions of kilometres across, making comets such as Hale-Bopp visible to the unaided eye. Until recently, the chemical composition of the nucleus was inferred from dissociation products in the visible spectra of comets. The advent of detectors working in new spectral domains — ultraviolet, radio and infrared — has directly confirmed the composition of cometary gases and extended the list to include more than 20 volatile species². One of the most important is CO, which in some comets is more than 10% relative to water.

Most of the time comets orbit far from the Sun, allowing them to remain virtually unchanged since the formation of the Solar System. The Sun's radiation is too weak to trigger the sublimation of water when comets are farther than about 4 AU, but some comets, such as Hale-Bopp, remain active far from the Sun (Fig. 1). Another, more volatile species must be responsible for activity at such distances. Carbon monoxide was

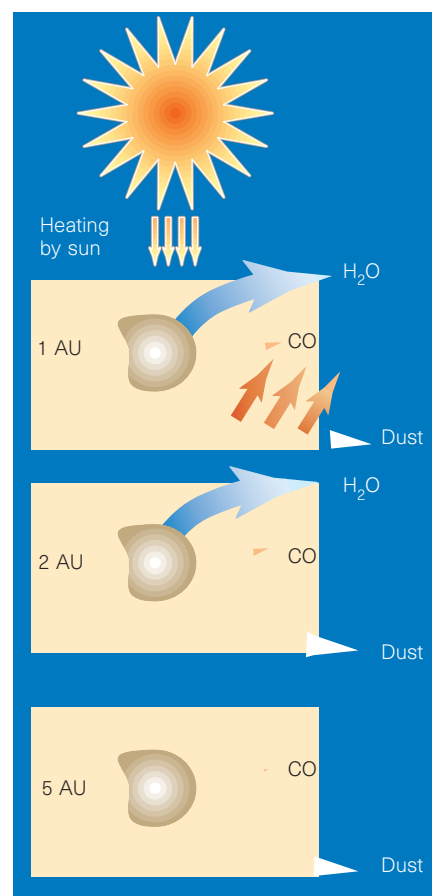


Figure 1 Cometary activity is governed by heat from the Sun, which vaporizes ices from the nucleus. At 5 AU (1 AU = 1 astronomical unit, the distance from Earth to Sun, or 150 million kilometres), water vaporization is inefficient, and CO accounts for cometary activity by dragging cometary dust particles from the ice matrix. At 2 AU, both H₂O and CO are driving cometary activity. At 1.5 AU or less, DiSanti *et al.*¹ have found that, in addition to the CO coming directly from the nucleus, CO is also released by a distributed source within the cometary atmosphere, which could be dust or an unidentified precursor. The production of CO from the nucleus and from the unidentified source is nearly equal.

found to be this volatile, first from radio observations of an active comet orbiting in a nearly circular orbit at about 6 AU, just beyond Jupiter³. For comet Hale-Bopp, which spends most of its time in the outer reaches of the Solar System, similar observations were able to detect CO when it was 7 AU from the Sun, beyond the orbit of Jupiter. Continuous monitoring of Hale-Bopp as it passed the Earth revealed a transition from CO-driven to water-driven activity as the comet approached the Sun⁴ (Fig. 1).

DiSanti *et al.*¹ used a long-slit spectrometer to observe the infrared CO distribution in one direction across the comet atmosphere. Far from the Sun, the CO distribution is found to be consistent with a direct release from the nucleus. But at 1.5 AU or less, a sec-

ondary source of CO, distributed in the 10,000-km-radius of the inner coma, is required to explain the spatial pattern of CO. Roughly equal amounts of CO are produced from the nucleus and from the extended source (Fig. 1). This means that CO measurements cannot be used to infer rates of mass loss from the nucleus, because a substantial fraction of the CO is being produced indirectly, perhaps from a precursor species in the coma. In addition to native water and CO, the amounts of carbon dioxide, methanol and methane found in comets are comparable to those existing in interstellar ices. This strengthens the link between comets and the primordial material from which the Solar System formed.

Spacecraft observations of comet Halley in 1986 revealed that CO and formaldehyde (H_2CO) were produced by an extended source^{5,6}. These one-off observations were made along the pathway of the Giotto space probe when it crossed the comet atmosphere, but they are difficult to interpret because of possible anisotropic gas outflow from the nucleus⁷. The secondary nature of the H_2CO source was subsequently confirmed by radio observations of this molecule, but the origin of the CO became a matter of debate.

What could be the nature of the unidentified extended source of these molecules? Icy grains dragged from the nucleus could be an important source of volatiles. But grains so close to the Sun are rapidly stripped of their frozen gases, and the volatiles they produce could not be distinguished from a nuclear source by ground-based observers. On the other hand, molecules such as carbon dioxide and methanol, which are important cometary volatiles, do produce CO when photodissociated by the Sun's ultraviolet radiation. But their photodissociation rate is too slow to account for the observed source of CO.

Another possibility would be the degradation of polymers or large organic molecules that may coat cometary grains. The presence of grains with organic crusts was suggested by mass spectroscopy of comet Halley's dust, which revealed a class of grains with a high content of CHON elements⁸. These were used to explain the distributed sources of CO and formaldehyde in comet Halley^{5,6}. Small molecules could be released from this material by sputtering following impacts of high-energy solar particles, by photolysis under the influence of solar ultraviolet radiation, or by pyrolysis when the grains are heated by the Sun. The observation that the distributed source of CO is triggered at 1.5–2 AU from the Sun¹ suggests that the last of these mechanisms could be at work. But, when trying to model these phenomena quantitatively, Greenberg and Li⁷ failed to explain the distributed source of CO in comet Halley, except by invoking very porous grains that could be overheated at high temperatures. Laboratory

simulations are urgently needed to investigate this problem⁹.

How could we unambiguously identify these distributed sources of volatiles in the coma? Perhaps by direct sampling of comets in future space missions. The Stardust spacecraft¹⁰, launched by NASA in February 1999, will bring back to Earth some samples of cometary dust particles, although we cannot be certain that their organic coatings will be preserved. The European Space Agency's Rosetta probe¹¹, to be launched in 2003, will orbit comet Wirtanen and send a lander to its surface in 2011–2013. The NASA project Deep Space 4/Chimpollion¹², also to be launched in 2003, will analyse material from the nucleus of comet Tempel 1 in 2005–2006. It is to be hoped that these missions will provide concrete evidence of the nature and composition of cometary material, and

reveal if and how cometary grains could trap volatiles in the form of ices and organic crusts. □

Jacques Crovisier is at the Observatoire de Paris, F-92195 Meudon, France.

e-mail: jacques.crovisier@obspm.fr

1. DiSanti, M. A. *et al.* *Nature* **399**, 662–665 (1999).
2. Crovisier, J. *Faraday Discuss.* **109**, 437–452 (1998).
3. Senay, M. & Jewitt, D. *Nature* **371**, 229–231 (1994).
4. Biver, N. *et al.* *Science* **275**, 1915–1918 (1997).
5. Eberhardt, P. *et al.* *Astron. Astrophys.* **187**, 481–484 (1987).
6. Meier, R., Eberhardt, P., Krankowsky, D. & Hodges, R. R. *Astron. Astrophys.* **277**, 677–690 (1993).
7. Greenberg, J. M. & Li, A. *Astron. Astrophys.* **332**, 374–384 (1998).
8. Krueger, F. R. & Kissel, J. *Naturwissenschaften* **74**, 312–316 (1987).
9. Cottin, H., Gazeau, M. C. & Raulin, F. *Planet. Space Sci.* (in the press).
10. <http://stardust.jpl.nasa.gov>
11. <http://sci.esa.int/missions/rosetta>
12. <http://nmp.jpl.nasa.gov/st4>

Neurobiology

Turning a corner in vision research

Ulf Eysel

The first step in processing of visual signals is carried out by a region of the brain known as the primary visual cortex. Although this region is not as highly specialized as some of the areas that mediate subsequent stages of processing, many basic properties of image analysis are generated here. For example, the primary visual cortex mediates orientation specificity — the preferential response of particular neurons to lines orientated in a narrow range of angles. Neurons that prefer the same orientation are close neighbours, forming vertical orienta-

tion columns that span the whole depth of the cortex. On page 655 of this issue, Das and Gilbert¹ report that these columns are used to analyse angular visual features in the context of a scene at this early stage of processing.

Columns of neurons with a preference for different orientations are distributed across the cortex in a characteristic orientation map² (Fig. 1). Columns with the same orientation form continuous bands or patches, and 'pinwheels'³ are formed where the different orientations meet. On average, a region

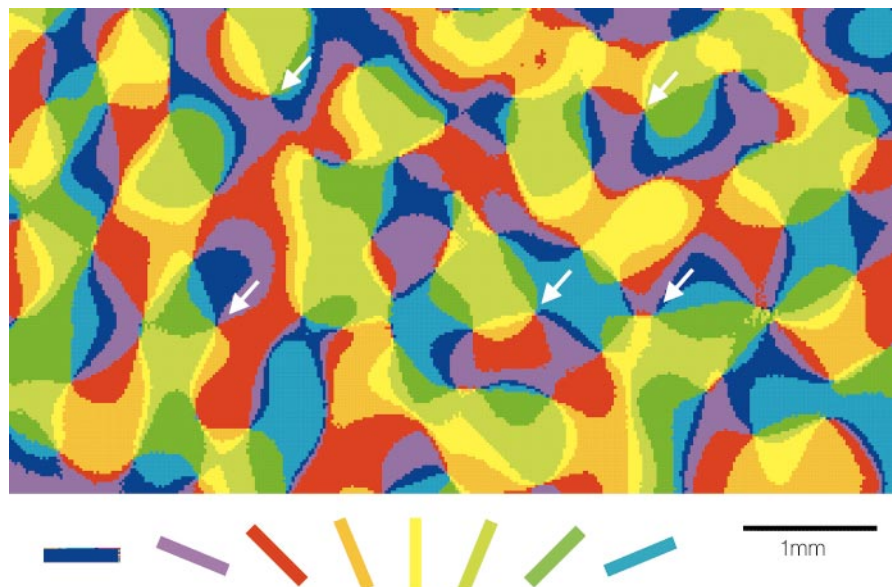


Figure 1 Map of preferred orientations across the surface of the primary visual cortex. Colour coding shows regions (orientation domains) with a preference for lines orientated at different angles — such as horizontal or vertical. Note the 'pinwheel' patterns (some marked with arrows), and the many regions where, at distances of about 0.5 mm, columns with very different orientation preferences are found. (Courtesy of Z. F. Kisvárdy.)

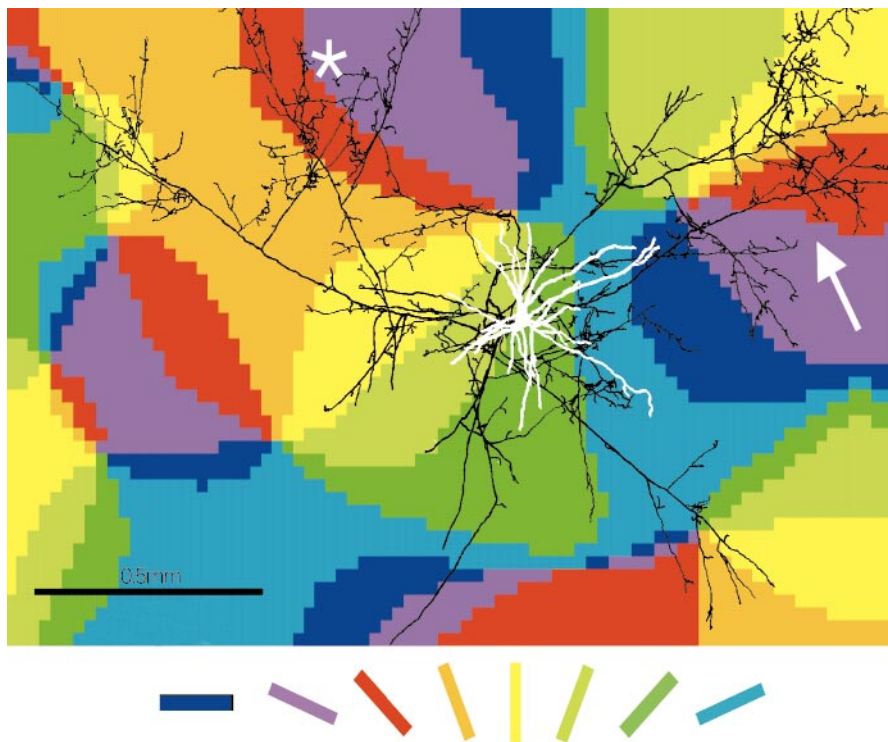


Figure 2 Orientation map across the primary visual cortex with superimposed inhibitory neuron. Das and Gilbert¹ have shown that inhibitory horizontal connections between perpendicular orientations across a pinwheel can be used to analyse angular visual features, such as corners or T junctions. An inhibitory layer 2/3 large basket cell is shown, with soma (cell body) and dendrites in white, and the axonal distribution (black). (Modified from ref. 6.) The axons innervate regions of similar (green) as well as perpendicular (red) orientation preferences. Over distances of up to 1 mm, this inhibitory cell can supply cross-orientation inhibition to regions with complementary colours across a pinwheel (arrow), as well as across different orientation domains (asterisk) not including a pinwheel.

of just 1 mm² on the surface of the cortex will contain all the possible orientation preferences, and, accordingly, can analyse orientation for one small area in the visual field. This topographical arrangement allows closely spaced objects with different orientations to interact. But it also means that a continuous line across the whole visual field would be cortically depicted in a patchy, discontinuous fashion. How can the spatially separated elements be bound together functionally?

The answer is that horizontal connections, varying in length from a few micrometres to 5–6 mm, allow local and remote interactions. They provide many excitatory and inhibitory connections between columns with both similar and dissimilar orientation preferences. Although the connections between similar columns could be useful in binding the distributed parts of a continuous contour, neurobiologists have long debated the functional significance of connections between drastically dissimilar columns. Now, Das and Gilbert¹ elegantly show that these connections have a suppressive (that is, inhibitory) function, and that, at this early stage of visual cortical processing, they could be used to analyse angular visual features such as corners or T junctions.

Forty years ago, Hubel and Wiesel⁴ char-

acterized the area of the visual field that can elicit responses from a visual cortical neuron as that cell's 'receptive field'. They also proposed that orientation specificity comes from aligned, convergent, excitatory subcortical inputs. Since then, many details about connectivity and processing within the cortex have been added. For example, if cortical inhibition is blocked, this can alter the orientation tuning of cells in the primary visual cortex⁵. One-fifth of all cortical neurons are now known to belong to one of seven distinct classes of inhibitory cell⁶. Unlike the excitatory cells, which preferentially connect to cells with similar properties⁷, the long-ranging axons of inhibitory cells connect equally to cells with similar and different orientation preferences, and even to cells with preferences at a right angle⁸ (Fig. 2).

Das and Gilbert¹ present new evidence for the functional significance of these cortical inhibitory interactions. The authors combined the optical imaging of intrinsic signals² (to obtain a functional map of orientation preferences across the cortical surface) with microelectrode recordings from pairs of cells at variable distances (representing non-overlapping parts of a visual scene with a preference for lines of very different orientations). They show a suppres-



100 YEARS AGO

In connection with the letter of "F.G." in *NATURE* of June 8 [see *Nature* 399, 525; 1999], on the strawberry cure of gout, I may mention that last year, when strawberries were so plentiful in England, a lady residing in Kent, who had formerly spent several years in Ceylon, where she had suffered from the wasting and often fatal complaint known as "Ceylon sore mouth" (the chief symptom of which is ulceration of the mucous membrane of the digestive organs), having had a return of the malady, and being unwilling to go abroad to undergo the "grape cure," conceived the happy idea to try strawberries instead, confining her diet to several pounds of these a day with plenty of milk. The remedy was so effectual that after a few weeks she was entirely cured of her malady, and had grown stout and well again.

Mr. G. Clarke Nuttall contributes to the current number of *The Contemporary Review* a popular account of the dependence of the flavour of tobacco upon the activity of bacteria during that important stage ... known as fermentation. Interesting reference is made to the work of Suchsland, who examined the germs which he found in the fermenting heaps of the finest West Indian tobacco. This German bacteriologist isolated and cultivated these bacteria, and then introduced some into quantities of inferior German tobacco, which was subsequently transformed so that connoisseurs could not distinguish it from the finest brands of tobacco.

From *Nature* 15 June 1899.

50 YEARS AGO

The University of Edlbach was founded by French prisoners-of-war in Oflag XVII A (1940–45). Not content with lectures alone, the geologists made a thorough investigation of the area – only 400 metres square – enclosed within the barbed wire. No stone was left unturned, and trenches and secret tunnels provided many critical exposures. A microscope was constructed in the camp and equipped with polarizers improvised from piled cover glasses. Thin sections were mounted with a mixture of violin wax and edible fat. Only the determination of certain untwinned feldspars remained to be completed on the return to France.

From *Nature* 18 June 1949.

sive effect of horizontal connections that decays with distance. This finding agrees with the authors' own model — that a cell's dendritic arbor overlaps with nearby cortical columns — as well as with the functional topography of laterally projecting inhibitory neurons⁸.

The observed suppression marks angular discontinuities at the level of the primary visual cortex. This information is then available to the higher brain areas that are specialized for more complex and elaborate analyses. Das and Gilbert's finding adds another side to the modern view that the response field of a neuron in the primary visual cortex reflects the functional state of an extended cortical network. Excitation and inhibition subtly interact in analysing an object centred on the 'classical' receptive field, and this analysis is modulated by contextual information from an extended surrounding^{9–11}.

Cells with drastically different orientation specificities are effectively connected by inhibitory circuits over distances of up to 800 μm (ref. 8). But the functional significance of these connections seems to differ depending on whether non-overlapping or overlapping receptive fields are involved. Cells at different sides of a pinwheel (Fig. 1) represent different parts of the visual space¹¹ — that is, they have non-overlapping receptive fields — so inhibitory horizontal connections provide contextual information about the visual scene¹. But cells located at similar distances, yet not separated by a pinwheel, have overlapping receptive fields. In that case, the inhibitory circuitry strongly suppresses the cell's response to any lines at the centre of the receptive field that have different orientations to the preferred one. The result is a sharper orientation tuning¹².

This specific effect was demonstrated by Crook *et al.*¹³ when columns with an orientation specificity at right angles to that of recorded cells were reversibly inactivated at a distance of 500–700 μm . This inactivation released the recorded cells from 'cross-orientation' suppression, meaning that they could respond to the orientations represented in the inactivated column¹². It seems, then, that the continuously debated cross-orientation inhibition¹⁴ can contribute functionally to different computations within the cortex. The same inhibitory cortical circuitry can probably sharpen orientation tuning locally¹², and provide higher-level contextual information from spatially distinct regions¹. This suggests that the same structures in the cortex can be functionally exploited for very different tasks — here, in analysing distinct features depending on the position of a neuron within the orientation map of the primary visual cortex. □

Ulf Eysel is in the Department of Neurophysiology, Medical School, Ruhr University Bochum, 44780 Bochum, Germany.

e-mail: eyssel@neurophysiologie.uni-bochum.de

1. Das, A. & Gilbert, C. D. *Nature* **399**, 655–661 (1999).
2. Grinvald, A. *et al.* *Nature* **324**, 361–364 (1986).
3. Bonhoeffer, T. & Grinvald, A. *Nature* **353**, 429–431 (1991).
4. Hubel, D. A. & Wiesel, T. N. *J. Physiol.* **148**, 574–591 (1959).
5. Sillito, A. M. *J. Physiol.* **250**, 305–329 (1975).
6. Somogyi, P. *Soc. Neurosci. Abstr.* **12**, 583 (1986).
7. Gilbert, C. D. & Wiesel, T. N. *J. Neurosci.* **9**, 2432–2442 (1989).

8. Kisvárdy, Z. F. *et al.* *Eur. J. Neurosci.* **6**, 1619–1632 (1994).
9. Gilbert, C. D. *Neuron* **9**, 1–13 (1992).
10. Sillito, A. M. *et al.* *Nature* **378**, 492–496 (1995).
11. Bringuier, V. *et al.* *Science* **29**, 695–699 (1999).
12. Das, A. & Gilbert, C. D. *Nature* **387**, 594–598 (1997).
13. Crook, J. M. *et al.* *Eur. J. Neurosci.* **10**, 2056–2075 (1998).
14. Vidyasagar, T. R. *et al.* *Trends Neurosci.* **19**, 272–277 (1996).

Carbohydrate chemistry

Sugars out in the open

Ole Hindsgaul

Structural knowledge of naturally occurring, biologically important molecules is at the very foundation of modern biochemistry and biology. When an unprecedented and unanticipated type of molecular structure is suddenly discovered, one therefore imagines that either 'the microscope has got bigger', or that someone was lucky and stumbled onto something, or paid exceptional attention to detail. In the case of Vinogradov and Bock — who, in *Angewandte Chemie International Edition* (**38**, 671–674; 1999) report the discovery of a new type of sugar–sugar linkage in bacterial polysaccharides — all three events occurred simultaneously.

The procedures for determining the primary sequences of DNA and proteins or peptides are well known from textbooks, and commercial instruments are available that perform the required chemistry and analysis. This is not so in the case of oligosaccharides and polysaccharides. Polynucleic acids (such as DNA and RNA) and proteins are assembled from a limited number of monomers, whereas hundreds are available for building polysaccharides. Further, carbohydrate poly-

mers do not have to be linear but can be branched, and the joining of sugar units together in glycosidic linkages involves the creation of a new stereochemical centre at the point of attachment.

Free sugars are known to exist predominantly in interconverting cyclic forms; the stable six-membered rings are called pyranoses, and the less stable five-membered rings are known as furanoses (Fig. 1). The carbon C-1 is asymmetric, so there are two possible configurations for each ring depending on whether the OH group is below (α -configuration) or above (β -configuration) the plane of the ring. For almost all sugars in solution, much less than 1% of the open-chain aldehyde form (that which contains a carbonyl group — a carbon and oxygen connected by a double bond) is present at equilibrium (Fig. 1). It is therefore not surprising that the biosynthesis of oligosaccharides and polysaccharides involves enzymes that act on the cyclic forms of sugars.

The generally accepted and well-supported mechanism for the synthesis of all polysaccharides is summarized in Fig. 2

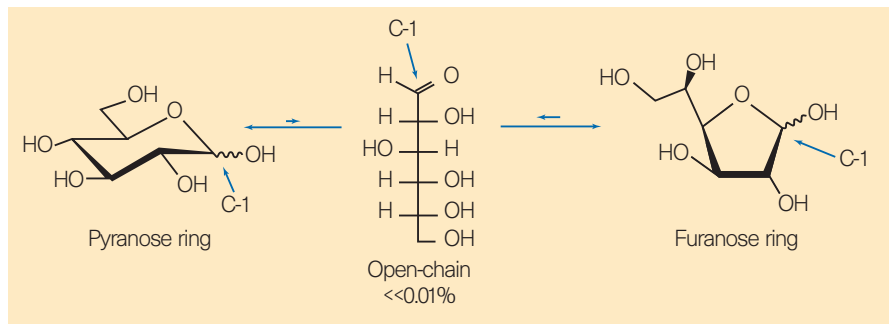


Figure 1 The composition of free sugars in solution.

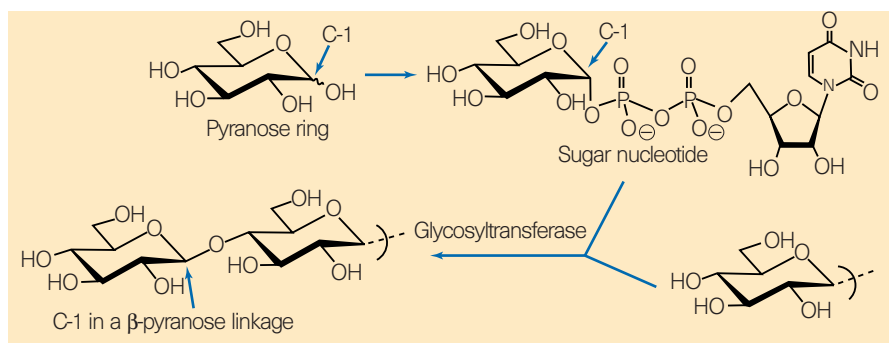


Figure 2 Biosynthetic process for joining together two sugars in a glycosidic linkage.

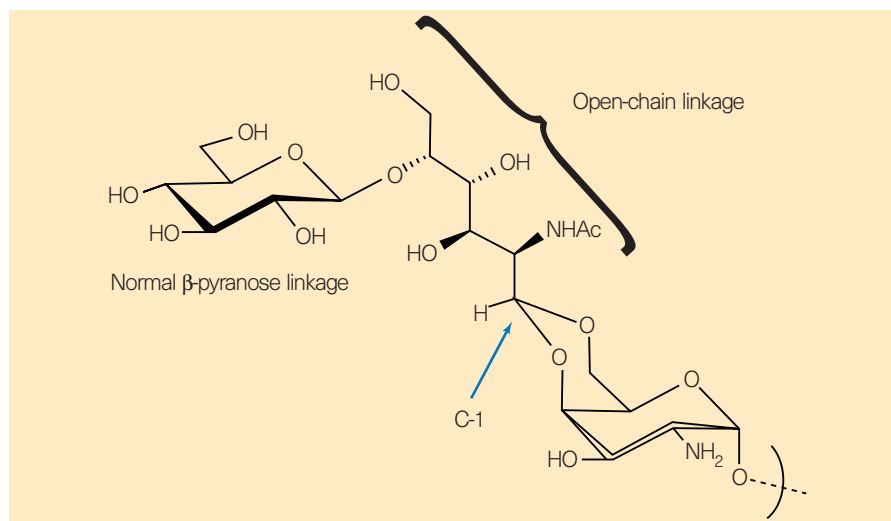


Figure 3 A new type of glycosidic linkage. The 'open-chain' linkage was discovered by Vinogradov and Bock¹ in the cell wall of *Proteus* bacteria. NHAc denotes an N-acetyl group.

(illustrated by the simple monosaccharide glucose). The cyclic pyranose form (or alternatively the cyclic furanose form) is phosphorylated at the C-1 hydroxyl group and ultimately converted to uridine diphosphate glucose, a high-energy sugar nucleotide. A glycosyltransferase enzyme then stereospecifically joins the cyclic pyranose to another sugar residue, for example in a β -pyranose linkage shown in Fig. 2. Repetition of this sequence using different sugar nucleotides and glycosyltransferases ultimately builds up the polysaccharides found in animals, bacteria and plants.

Vinogradov and Bock were fortunate to be at the right place at the right time. They were involved in a presumably routine study on the structure of the cell-wall polysaccharides of a poorly characterized *Proteus* microorganism. Many such structural studies have been reported for other bacteria, fuelled by both the scientific challenge of tackling complexity and by the hope of generating polysaccharide-based vaccines. On this occasion the authors had access to state-of-the-art instrumentation (high-field nuclear magnetic resonance) and they paid meticulous attention to detail. They concluded that their structural data were inconsistent with all known sugar–sugar linkages. They had discovered the first example of an open-chain glycosidic linkage (Fig. 3).

With the benefit of hindsight, the existence of open-chain glycosidic linkages involving sugar C-1 aldehydes makes perfect chemical sense. After all, aldehydes, such as benzaldehyde, have been extensively used as protecting groups in sugar chemistry. Moreover, pyruvate acetals (ketone compounds containing a carbonyl group that have reacted with sugar diols) are known to exist in microorganisms. But the reason open-chain glycosidic linkages have never even been considered is undoubtedly because we were all taught that sugars in nature are cyclic.

The bacterial polysaccharides found to contain these new linkages may or may not have important biological functions. They

Developmental neurobiology

Decoding the Reelin signal

Isabelle Bar and André M. Goffinet

Back in 1995, the *reeler* gene was cloned and its protein product was christened Reelin¹. Since then, the Reelin signalling pathway — which is involved in brain development — has been progressively unravelled. The latest, most unexpected, twist is now reported in *Cell*, where Trommsdorff *et al.*² show that mice lacking both the very low density lipoprotein (VLDL) receptor and apolipoprotein E receptor-2 (ApoE-R2) genes have the same characteristics as mice that lack Reelin.

During development, neurons generated along the cerebral ventricles migrate through the tissue, then settle into rigorously defined patterns before deploying their dendrites and axons to connect with other cells. Particularly exquisite patterns are formed in the cerebellum, inferior olive, hippocampus and, most of all, in the six-layered cerebral cortex and its precursor, the embryonic cortical plate (Fig. 1a, overleaf). But these patterns are disturbed in *reeler* mutant mice, which have been studied as a model of abnormal brain development for 50 years (reviewed in ref. 3).

Reelin is a large glycoprotein¹ found in the extracellular matrix that is secreted by, and surrounds, cells. Reelin is secreted by a variety of neurons, most notably by Cajal–Retzius cells in the cortical marginal zone, and by cerebellar granule cells. But it is not expressed by those neurons that show disrupted patterns (the *reeler* phenotype), such as cortical-plate neurons or cerebellar Purkinje cells.

may make an important contribution to the stability of the cell walls, or they could be important cell-wall antigens. They could even be virulence factors, or they could simply be unnecessary evolutionary baggage. But what is far more intriguing is that their biosynthesis requires the existence of a brand new class of enzymes, because the known biosynthetic pathways for polysaccharides use cyclic sugar nucleotides. It is tempting to speculate that the biosynthesis of open-chain glycosidic linkages may in fact involve the addition of a cyclic sugar followed by the action of a glycosidase cleaving enzyme, whose mechanism involves an open-chain intermediate. Such intermediates have been hypothesized, but never identified. It will be interesting to see just how common the open-chain glycosidic linkage is, now that we know that it exists and therefore know how to look for it. □

Ole Hindsgaul is in the Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada.

e-mail: ole.hindsgaul@ualberta.ca

This suggests that Reelin acts at a distance on target cells, via the extracellular matrix. Two years ago, it was shown that the *reeler* phenotype can also result from mutations of mouse Disabled-1 (Dab1). This phosphotyrosine protein, which resides in the cytoplasm, is thought to interact with tyrosine kinases of the Src and Abl families^{4–6}. Unlike Reelin, Dab1 is expressed in the neurons that manifest the *reeler* phenotype, suggesting that it is a key element of the response to a Reelin signal. Indeed, phosphorylation of Dab1 is decreased in Reelin-deficient embryos, but increased when Reelin is added to neuronal cultures⁷. Until now, the molecules in the cell membrane that sense the presence of Reelin in the extracellular matrix and relay this signal to the cell via Dab1 have been a mystery.

Enter Trommsdorff and colleagues². Mice with single mutations in the VLDL receptor gene have some abnormalities in neuronal patterning, and single mutants of the ApoE-R2 gene show subtle alterations in cortical development. But the double mutants develop a drastic malformation that is identical to the *reeler* phenotype, suggesting that the functions of the VLDL receptor and ApoE-R2 overlap. Both proteins are expressed in cells that contain Dab1 — cortical-plate neurons and Purkinje cells, for example — although ApoE-R2 is also found deeper in the tissue. The VLDL receptor and ApoE-R2 proteins belong to a small family that also includes the low density lipoprotein receptor of Nobel fame, megalin and the

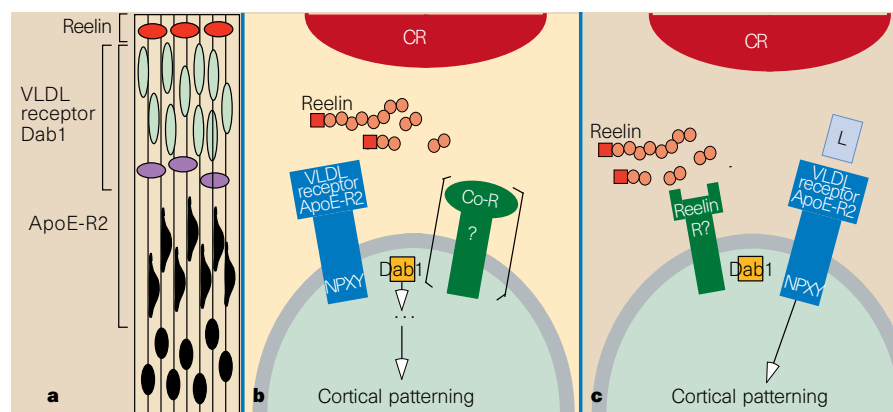


Figure 1 Lipoprotein receptors and the Reelin signalling pathway. **a**, In the developing cortex, neuron precursors (black ovals) give rise to immature, elongated neurons that migrate (black) to settle in the cortical plate (green). This area is bordered externally by Cajal–Retzius cells (CR; red), which secrete Reelin, and internally by the subplate (pink). Patterning of the cortical plate depends not only on Reelin, but also on proteins expressed in the cortical-plate cells — Dab1 and, according to Trommsdorff *et al.*², the VLDL receptor and the ApoE-R2. **b**, The VLDL receptor and ApoE-R2 could react to the presence of Reelin or Reelin fragments by docking Dab1 through their NPXY cytoplasmic sequence. A co-receptor (Co-R) to Reelin could also be present. **c**, Reelin could act on an unidentified Reelin receptor and, via Dab1, recruit the VLDL receptor and ApoE-R2. Then, the VLDL receptor and ApoE-R2, together with an associated putative extracellular ligand (L), could direct cortical patterning.

LDLR-related protein. Apparently, the other three members of this family do not share the effects of the VLDL receptor and ApoE-R2 on brain development.

The best-known ligand for the VLDL receptor and ApoE-R2 is apolipoprotein E (ApoE). Although this protein is synthesized in the brain — probably in glial cells — it is unlikely to be critical in the Reelin pathway, because mice that lack ApoE develop normally. As well as ApoE, lipoprotein receptors can bind thrombospondin-1, lipoprotein lipase, the serine protease urokinase-type plasminogen activator and probably others. The ligands implicated in patterning the embryonic brain have not, however, been defined thus far.

The fascinating finding that Reelin, Dab1 and the VLDL receptor/ApoE-R2 participate in the same genetic pathway fits nicely with previous observations. For instance, Dab1 interacts with the cytoplasmic tail of lipoprotein receptors and binds directly to the NPXY sequence, which has been implicated in lipoprotein-receptor-mediated cellular uptake^{8,9}. The Dab1 protein also docks to the cytoplasmic signalling protein Ship, to some phosphatidylinositols, as well as to the β -amyloid precursor protein and related proteins (ref. 9; R. Homayouni and T. Curran, personal communication). The β -amyloid precursor protein is cleaved to form the amyloid deposits in Alzheimer's disease. Moreover, the ApoE4 allele is associated with an increased predisposition to this disease. So, these observations might have pathophysiological implications.

Although several pieces of the puzzle are still missing, Trommsdorff and colleagues' data can be interpreted in at least two ways. The most parsimonious explanation (Fig. 1b) is that the VLDL receptor and ApoE-R2 are

receptors to Reelin or fragments of this protein¹⁰. By docking Dab1, these receptors then modulate associated tyrosine kinase(s), and instruct the patterning of neurons at the end of migration. A variant of this explanation is that the VLDL receptor and ApoE-R2 could bring Dab1 to a Reelin co-receptor that has an associated tyrosine kinase activity. The main difficulty with this model is that, as far as we know, nobody has yet shown a physical association between the VLDL receptor/ApoE-R2 and Reelin or Reelin fragments.

A different view (Fig. 1c) is that the VLDL receptor/ApoE-R2 could act downstream of Reelin and Dab1. Lipoprotein receptors are involved in cellular uptake so they could, for example, modify membrane trafficking, thereby shaping the normal characteristics of postmigratory neurons. As yet, though, the putative Reelin receptor, its interaction with Dab1, and the ligand(s) for the lipoprotein receptor remain to be characterized. Whichever model (if either) is correct, both point to the necessity to identify and define putative Reelin receptors. This task will not be any the easier if the processing of Reelin into fragments, which has now been confirmed by several groups, is functionally important. □

Isabelle Bar and André M. Goffinet are in the Neurobiology Unit, University of Namur Medical School, 61 rue de Bruxelles, B5000 Namur, Belgium. e-mail: Andre.Goffinet@fundp.ac.be

1. D'Arcangelo, G. *et al.* *Nature* **374**, 719–723 (1995).
2. Trommsdorff, M. *et al.* *Cell* **97**, 689–701 (1999).
3. Lambert de Rouvroit, C. & Goffinet, A. M. *Adv. Anat. Embryol. Cell. Biol.* **150**, 1–106 (1998).
4. Howell, B. W. *et al.* *Nature* **389**, 733–737 (1997).
5. Sheldon, M. *et al.* *Nature* **389**, 730–733 (1997).
6. Ware, M. L. *et al.* *Neuron* **19**, 239–249 (1997).
7. Howell, B. W. *et al.* *Genes Dev.* **13**, 643–648 (1999).
8. Trommsdorff, M. *et al.* *J. Biol. Chem.* **273**, 33556–33560 (1998).
9. Howell, B. W., *et al.* *Mol. Cell. Biol.* (in the press).
10. Lambert de Rouvroit, C. *et al.* *Exp. Neurol.* **156**, 214–217 (1999).

Daedalus

Total digital recall

How we store our memories is still a mystery. The secret seems to lie in the synapses where a dendrite from one brain cell touches the axon of another. But the information may be recorded physically, via the growth of new synapses; or chemically, via the synthesis of special encoding substances. These alter the chance that the synapse will convey a nervous impulse from one neuron to the other, or the chance that the receiving neuron will fire under this stimulation.

To settle the matter by experimenting on living brains is daunting, and maybe unethical as well. So Daedalus plans to study the information content of dead ones. Already organizations exist which will freeze your corpse in liquid nitrogen, and store it until science is sufficiently advanced to thaw you out and restore you to life. If you cannot afford a whole-body cryogenic crypt, you can merely have your head stored instead. In due course, science will then have to give you a new body. Alternatively, says Daedalus, it might be possible to read all the information in your dead and frozen brain. The totality of your experience would thus be captured for posterity — a rather second-class form of immortality, but better than nothing.

DREADCO physiologists are now working out how to do it. They propose sectioning the frozen brain into micron-thin slices with a cryogenic microtome. Each slice will be scanned by electron microscope, and recorded as a digitized image. The resulting 100,000 images will be stored in a terabyte memory, and used to reconstruct the complete synaptic 'wiring diagram' of the brain. Its chemical map will be harder to recover. Each slice will have to be treated with fluorescent immunoassay reagents binding to the crucial encoding substances, so as to record their distribution in the slice.

Daedalus cheerfully admits that the resulting vast mass of data will be utterly incomprehensible. But, as with the steadily growing incomprehensible data of the human-genome project, we must hope that one day it will all make sense. In due course science will advance enough to interpret your stored brain map — when your heroic sacrifice will finally bear fruit. Every detail of your life will be recovered, as a fascinating historical record. Secrets you took to the grave will be revealed at last. And if the data can by then be inserted back into a living brain, you might live again as a parasitic consciousness within the mind of someone else.

David Jones

Visual kin recognition in chimpanzees

The ability to distinguish between members of one's own species has greatly assisted the evolution of sociality in mammals, leading to individualized relationships and cooperative networks. Because kin selection is important for the evolution of complex societies, other advantages must derive from the ability to distinguish kin from non-kin¹. Taking advantage of the chimpanzee's face-recognition abilities and computer skills^{2,3}, we presented five subjects with the task of matching digitized portraits of unfamiliar females with their offspring. We find that chimpanzees can match the faces of mothers and sons, but not mothers and daughters, providing evidence for a mechanism of kin recognition in primates that is independent of previous experience with the individuals in question.

Two separate kin-recognition mechanisms have been proposed: early familiarity

and phenotypic matching^{4,5}. Phenotypic matching by means of non-visual cues has been demonstrated in several non-primates, but not in primates^{6–8}. Instead, kin recognition in primates seems to depend on previous experience, for example the ability of macaques to match photographs of familiar mothers and offspring⁹. But phenotypic matching would have the advantage of providing information about the kin status of individuals unknown to the observer, like the facial similarities that can be seen in human family albums.

Chimpanzees (*Pan troglodytes*) form loosely organized fission–fusion communities in the wild, in which even closely related individuals spend considerable time apart. Association is therefore not as reliable an index of kinship as in the best-studied monkey species, in which kin preferences seem to be socially learned¹⁰.

We examined the chimpanzees' ability to recognize facial similarities in black-and-white portraits of unfamiliar conspecifics on a computer screen by using four types of discrimination. Individual recognition (IR) trials presented two different photographs of the same individual. Mother–offspring trials presented photographs of mothers and daughters (MD) or mothers and sons (MS) (Fig. 1). Finally, unrelated control (UC) trials presented photographs of individuals unrelated through maternal or paternal lines. We considered that subjects would perform best on the IR trials, and better on mother–offspring trials than on UC trials. Test subjects had never had any contact with the stimulus individuals, most of whom lived at distant locations. All offspring photographs depicted adults or subadults (more than seven years old).

Confirming earlier evidence of individual recognition², subjects did better on IR trials than on every other category (IR versus MS, $t(4) = 2.48$, $P < 0.05$; IR versus MD, $t(4) = 5.87$, $P < 0.005$; IR versus UC, $t(4) = 4.46$, $P < 0.01$). Subjects also did better on MS than on UC trials ($t(4) = 2.80$, $P < 0.025$), but unexpectedly no significant difference was found for performance on MD compared with UC trials. This means that subjects saw no more facial similarity between mothers and daughters than between unrelated individuals. Within the mother–offspring category, performance was significantly better for MS than for MD trials ($t(4) = 3.58$, $P < 0.05$). There was no overlap in performance on MD and MS pairs; the highest individual performance on MD trials (61.1%) was below the lowest performance on MS trials (64.5%). The mean performance on all stimulus categories, including the mean performance

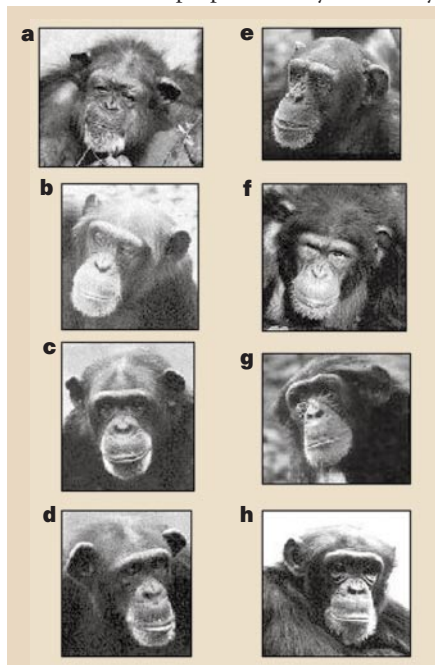


Figure 1 Mother–offspring pairs, showing two examples of each category: mother–daughter and mother–son. Left, mothers; right, offspring. The correct combinations here are **a** with **g** (daughter), **b** with **f** (son), **c** with **h** (son), and **d** with **e** (daughter). Stimuli were presented using a computerized testing model and a matching-to-sample format, which relied on the apes' skill to indicate their choice with a joystick-controlled cursor on the screen. In each trial, a single portrait of an adult female was presented as the sample, after which subjects were required to choose between two comparison portraits that matched the sample depending on whether it was a different portrait of the same female (IR trials), an offspring of this female (MD and MS trials), or an arbitrarily designated non-relative (UC trials). Further details of the test paradigm are available from the authors and in refs 2, 3.

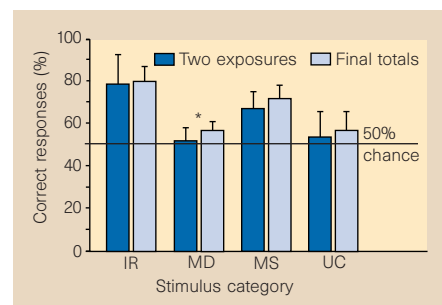


Figure 2 Mean performance for each stimulus category. Mean percentage ($\pm$ s.d.) of correct responses is compared for each of the four discriminations (IR, individual recognition; MD, mother–daughter; MS, mother–son; UC, unrelated controls) with performance on the first day of testing. Whereas cumulative totals concerned between 604 and 648 trials per subject, the first two exposures to IR pairs totalled 24 trials, with 20 trials each of MD, MS and UC pairs. The asterisk marks the category in which performance was significantly better at the end of testing than on the first day ($P < 0.05$, two-tailed). This learning effect was evident only for mother–daughter pairs.

for the first testing session, is shown in Fig. 2. The superior performance on mother–son as opposed to mother–daughter pairs was evident from the onset of testing.

These findings indicate that chimpanzees perceive similarities in the faces of related but unfamiliar individuals, providing evidence of visual kin recognition at a purely phenotypic level. But this ability seems to be limited to recognizing mothers and sons. Perhaps facial similarities to do with sons, or males in general, are more noticeable to chimpanzees. Such an attentional bias would make sense from an evolutionary perspective in view of the chimpanzee's male philopatric society and the tendency towards 'political' alliances in which males incur great risk on behalf of other males¹¹. Phenotypic matching might assist the recognition of subsets of related males who tend to support each other. The capacity might also have a role in inbreeding avoidance: a migrating young female should probably not settle in a community in which many males look like her mother, as these males might be related to her.

Lisa A. Parr, Frans B. M. de Waal

Living Links, Yerkes Regional Primate Research Center, and Department of Psychology, Emory University, Atlanta, Georgia 30329, USA
e-mail: parr@rmy.emory.edu

- Gouzoules, S. & Gouzoules, H. in *Primate Societies* (eds Smuts, B. B. et al.) 299–305 (Univ. Chicago Press, 1987).
- Parr, L. A., Winslow, J. T., Hopkins, W. D. & de Waal, F.B.M. *J. Comp. Psychol.* (in the press).
- Parr, L. A., Dove, T. A. & Hopkins, W. D. *J. Cog. Neurosci.* **10**, 615–622 (1998).
- Holmes, W. G. & Sherman, P. W. *Am. Sci.* **71**, 46–55 (1983).

5. Lacy, R. C. & Sherman, P. W. *Am. Nat.* **121**, 489–512 (1983).
6. Erhart, E. M., Coelho, A. M. Jr & Bramblett, C. A. *Am. J. Primatol.* **43**, 147–157 (1997).
7. Fredrickson, W. T. & Sackett, G. P. *J. Comp. Psychol.* **96**, 29–34 (1984).
8. Welker, C., Schwibbe, M. H., Schäfer-Witt, C. & Visalberghi, E. *Folia Primatol.* **49**, 216–221 (1987).
9. Dasser, V. *Anim. Behav.* **36**, 225–230 (1988).
10. de Waal, F. B. M. *J. Comp. Psychol.* **110**, 147–154 (1996).
11. de Waal, F. B. M. *Chimpanzee Politics: Power and Sex Among Apes* (Jonathan Cape, London, 1982).

The oldest fossil ascomycetes

Ascomycetes are the largest group of true fungi, and characteristically produce their sexual spores in a sac-like structure called the ascus. They include medicinal agents (such as ergot), plant pathogens (Dutch elm disease is caused by an ascomycete) and yeasts used in fermentation. We have found the oldest ascomycetous fungi with flask-shaped ascocarps in thin-section preparations of the Lower Devonian (400 million years old) Rhynie chert of Aberdeenshire, Scotland¹. This discovery has implications for dating the origin of this group of fungi, and underscores the diversity of

fungal–plant interactions early in the colonization of the land.

The fossils occur as closed fruiting bodies (perithecia) scattered just beneath the epidermis in the upright stems and rhizomes of the early land plant *Asteroxylon* (Fig. 1a); a few ascocarps have been identified on the scale leaves of this plant. Each globose perithecium (Fig. 1b) is approximately 400 µm in diameter and develops in the stomatal chamber beneath a pair of guard cells. At maturity, the perithecium has a slightly elongate neck through which the spores are released. The perithecium wall is constructed of two distinct layers of interwoven hyphae. Arising from the inner layer are numerous tightly packed, elongate asci (Fig. 1c), each up to 50 µm long. The ascus wall is thin and appears to consist of a single layer. Some asci have a slight invagination that encircles the tip of the ascus, suggesting the presence of some structural modification for spore release.

Arising from the same layer as the asci are delicate, thread-like structures. These may represent some form of sterile hairs (paraphyses) or be the remains of asci that have discharged their spores. Sterile hairs (periphyses) line the neck canal of each perithecium. Each ascus contains 16, and

perhaps up to 32, elongate ascospores (Fig. 1c). Each ascospore is approximately 5 µm long, and many are bicelled (Fig. 1d). In other regions along the stems, tufts of hyphae extend through the cuticle of *Asteroxylon* and produce chains of conidiospores (non-sexual spores) at their tips.

The fossil history of the Ascomycota is poorly understood². Although chains of asexual spores and perforate hyphae recovered from digested rock samples of Silurian age have been interpreted as ascomycetes³, they may be modern contaminants or remains of some non-fungal group⁴. Molecular clock estimates have continued to push back in geological time the divergence of major fungal groups, with the most recent estimates suggesting that the basidiomycetes and ascomycetes diverged from one another about 500 million years ago⁵.

Comparative gene sequence data have identified three groups of Ascomycota: Archiascomycetes (unicellular yeast-like forms), Hemiascomycetes (yeasts) and Euascomycetes (filamentous forms). Euascomycetes are a monophyletic group that includes the pyrenomycetes⁶, which have elongate asci that arise as a single layer within a perithecium, and ascospores that are sometimes forcibly discharged at maturity. The most basal members of the Euascomycete group have not been resolved.

The Rhynie chert ascomycete fossils contain characters of the sexual stage of the fungus (such as perithecium, asci and ascospores) that are morphologically identical to those found in modern pyrenomycetes. They therefore provide the oldest evidence by approximately 250 million years of a major lineage of Euascomycetes with perithecial ascocarps. This is important because in some molecular trees the pyrenomycetes appear to be relatively recent compared with other ascomycetes. Because the characters in the Rhynie chert fossils are so well preserved, they can help to establish the sequence of phylogenetic events within this group. They therefore provide a means of calibrating the molecular clock data sets that are used to infer the phylogeny of fungal groups.

T. N. Taylor*, H. Hass, H. Kerp†

*Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas 66045, USA

e-mail: ttaylor@falcon.cc.ukans.edu

†Geologische-Paläontologische Institut und Museum, Westfälische Wilhelms-Universität Münster, 4400 Münster, Germany

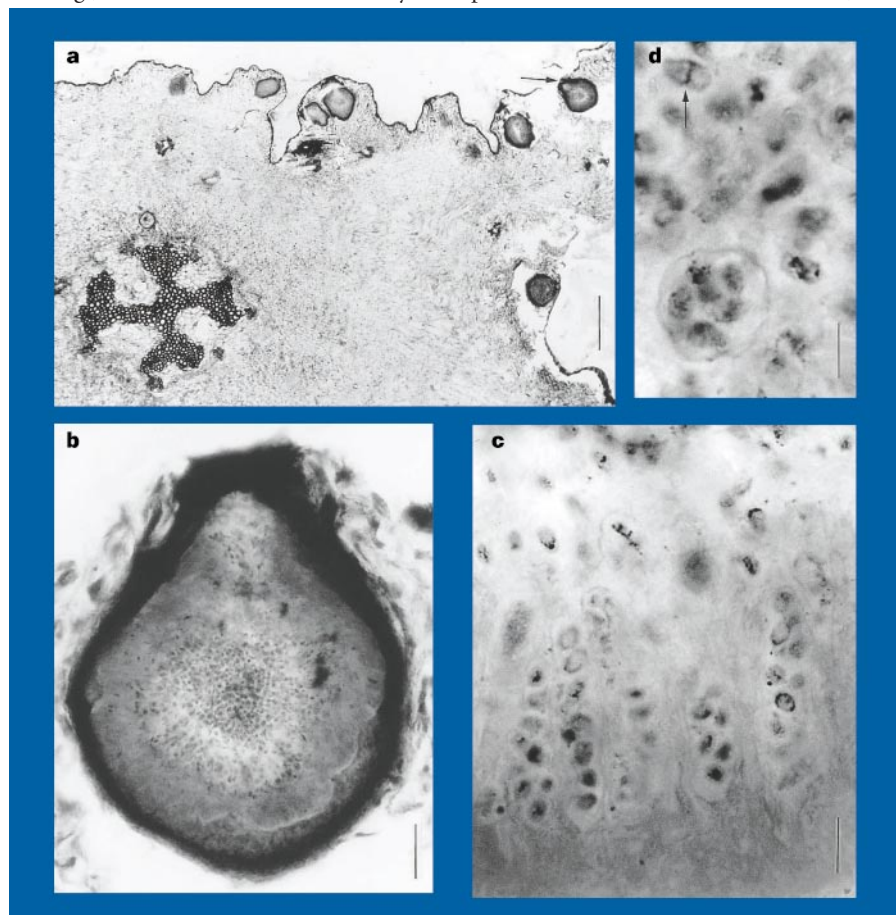


Figure 1 Lower Devonian ascomycete. **a**, Section of stem with ascocarps (arrow) in cortex. **b**, Perithecium in longitudinal section. **c**, Asci containing ascospores arising from inner wall of perithecium. **d**, Cross-section of ascus (lower left) with ascospores, some of which are bicelled (arrow). Scale bars: **a**, 500 µm; **b**, 50 µm; **c**, 10 µm; **d**, 20 µm.

1. Taylor, T. N., Hass, H. & Kerp, H. *Am. J. Bot.* **84**, 992–1004 (1997).
2. Taylor, T. N. in *Ascomycete Systematics* (ed. Hawksworth, D. L.) 167–173 (Plenum, New York, 1994).
3. Sherwood-Pike, M. A. & Gray, J. *Lethaia* **18**, 1–20 (1985).
4. Berbee, M. L. & Taylor, J. W. *Can. J. Bot.* **71**, 1114–1127 (1993).
5. Berbee, M. L. & Taylor, J. W. in *Mycota* (Springer, Berlin, in the press).
6. Alexopoulos, C. J., Mims, C. W. & Blackwell, M. *Introductory Mycology* (Wiley, New York, 1996).

Is there solar argon in the Earth's mantle?

Pepin¹ suggests, mainly on the basis of three studies^{2–4}, that solar argon, krypton and xenon might be present in the Earth's mantle. But I believe that the analytical evidence is still too weak to claim such a breakthrough in rare-gas geochemistry. Moreover, other studies^{5–8}, and the new data I report here, indicate that solar argon, krypton and xenon, if present at all, are only a minor constituent of mantle rare gases.

Neon isotopes in samples from the Earth's mantle clearly indicate a solar-like composition (see, for example, ref. 1) once the effect of atmospheric contamination is removed. Models that explain this feature also involve the presence of solar-like argon in the mantle, which should yield lower-than-air $^{38}\text{Ar}/^{36}\text{Ar}$ ratios.

Burnard *et al.*² took laser measurements of CO_2 , ^4He , ^{36}Ar and ^{40}Ar in a volatile-rich mid-ocean ridge basalt (MORB) glass (2IID43) and inferred a solar-like elemental abundance pattern for all rare-gas species in the mantle. However, Moreira *et al.*⁵ analysed all rare-gas isotopes in the same sample and showed that this abundance pattern is close to the atmospheric (planetary) curve, even when corrected for air contamination by extrapolation to 'pure' solar neon ($^{20}\text{Ne}/^{22}\text{Ne}=13.8$). Valbracht *et al.*³ found low $^{38}\text{Ar}/^{36}\text{Ar}$ ratios correlating with non-atmospheric $^{20}\text{Ne}/^{22}\text{Ne}$ and $^{40}\text{Ar}/^{36}\text{Ar}$ ratios in submarine basalts from the Loihi hotspot, and claimed the first but "still preliminary" evidence for solar argon in the mantle. However, if Pepin's range for solar argon (0.1724–0.1786)¹ is used, the uncontaminated $^{40}\text{Ar}/^{36}\text{Ar}$ ratios deduced from $^{20}\text{Ne}/^{22}\text{Ne}$ versus $^{40}\text{Ar}/^{36}\text{Ar}$ and $^{38}\text{Ar}/^{36}\text{Ar}$ versus $^{40}\text{Ar}/^{36}\text{Ar}$ systematics agree less well, which may weaken the evidence in ref. 3.

The significance of the low $^{38}\text{Ar}/^{36}\text{Ar}$ ratios found in 'plume on the ridge' samples⁴ is even less clear, because they do not correlate with elevated $^{40}\text{Ar}/^{36}\text{Ar}$: they are also observed for atmospheric $^{40}\text{Ar}/^{36}\text{Ar}$

ratios. I therefore agree with the authors of ref. 4 that their data are "compatible with the air ratio" and that any deviation observed is hardly significant (S. Niedermann, personal communication).

I am sceptical about this evidence when I look at the results of other more sensitive studies^{6,7} in which no significant deviation from the atmospheric $^{38}\text{Ar}/^{36}\text{Ar}$ ratio was observed that was accompanied by elevated $^{20}\text{Ne}/^{22}\text{Ne}$ or $^{40}\text{Ar}/^{36}\text{Ar}$ ratios, even though the technique used could have detected such variations if the effect in the Loihi data were real. Solar xenon also appears to be negligible in the MORB source⁸.

I have studied popping rock (2IID43) to look for solar argon in the MORB source (Table 1), taking care to avoid impurities, interference and mass-fractionation effects, which could have biased the data. All of our fractions except one reveal clearly non-atmospheric neon isotope ratios close to $^{20}\text{Ne}/^{22}\text{Ne}=12.5$ and $^{21}\text{Ne}/^{22}\text{Ne}=0.06$, and $^{40}\text{Ar}/^{36}\text{Ar}$ ratios correlate closely with them, as described⁵. But no correlation for the $^{38}\text{Ar}/^{36}\text{Ar}$ ratios (Fig. 1a) could be found.

A more rigorous argument can be derived from an argon three-isotope plot (Fig. 1b), in which two-component mixing gives a straight mixing line instead of hyperbolic curves in four isotope plots. A York fit⁹ reveals no slope significantly different from zero ($a=0.18822\pm0.00027$, $b=(0.752\pm1.726)\times10^{-8}$), which may indicate that there is no solar argon in the MORB source. Even taking an extreme choice of the confidence band ($a-2\sigma$, $b-2\sigma$), the $^{38}\text{Ar}/^{36}\text{Ar}$ ratios for reasonable estimates of MORB $^{40}\text{Ar}/^{36}\text{Ar}$ remain close to the atmospheric value. By using a $^{40}\text{Ar}/^{36}\text{Ar}$ value of 44,000 for uncontaminated MORB⁵, the corresponding $^{38}\text{Ar}/^{36}\text{Ar}$ is estimated to be greater than 0.1858 (likelihood > 97.5%). This value remains clearly distinct from solar $^{38}\text{Ar}/^{36}\text{Ar}$, even for higher MORB $^{40}\text{Ar}/^{36}\text{Ar}$. However, the best estimate derived from our data is atmospheric, no matter which mantle $^{40}\text{Ar}/^{36}\text{Ar}$ value is assumed.

I therefore think that solar argon, if present at all, represents only a negligible constituent of the upper-mantle non-radiogenic argon budget. This also seems to be true for

the more primitive source of the hotspots. Using the data and estimate for the uncontaminated $^{40}\text{Ar}/^{36}\text{Ar}$ of the source from ref. 7, a fairly similar lower limit for $^{38}\text{Ar}/^{36}\text{Ar}$ (0.1868) is obtained.

A possible explanation for these observations is that most ^{36}Ar and ^{38}Ar present in the mantle may have been re-injected from the atmospheric reservoir into the mantle, for example by subduction. Although subduction-related volcanism may be an effective barrier for rare gases¹⁰, even a small amount of atmospheric argon could dominate (pollute) the mantle budgets. If this model is true, then I can derive from my lower limit for mantle $^{38}\text{Ar}/^{36}\text{Ar}$ (calculated above) that at least 83% of these isotopes in the MORB source result from argon recycling by subduction. The first

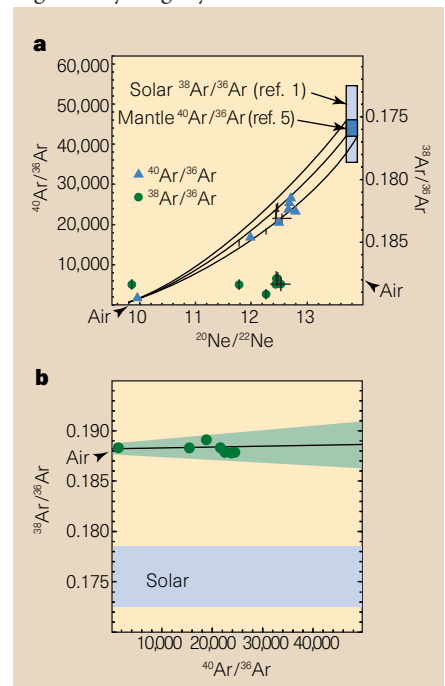


Figure 1 Isotope correlations. **a**, The $^{38}\text{Ar}/^{36}\text{Ar}$ versus $^{20}\text{Ne}/^{22}\text{Ne}$ diagram (green circles, right axis) shows no correlation, in contrast with the close correlation between $^{40}\text{Ar}/^{36}\text{Ar}$ and $^{20}\text{Ne}/^{22}\text{Ne}$ (blue triangles, left axis). Scaling was adjusted to reach an approximate overlap for the endmember compositions, atmospheric ($^{40}\text{Ar}/^{36}\text{Ar}=295.5$ and $^{38}\text{Ar}/^{36}\text{Ar}=0.188$) on the left and MORB source ($^{40}\text{Ar}/^{36}\text{Ar}=44,000$ (ref. 5) and suggested solar $^{38}\text{Ar}/^{36}\text{Ar}$ (ref. 1)) on the right. In such a four-isotope plot, two-component mixing gives a hyperbola. Such a hyperbola approaches a linear trend where the ratio of the normalization isotopes (here $^{20}\text{Ne}/^{36}\text{Ar}$) is equal in the two mixing reservoirs. This hyperbola is from ref. 5, in which $(^{20}\text{Ne}/^{36}\text{Ar})_{\text{uppermantle}}/(^{20}\text{Ne}/^{36}\text{Ar})_{\text{air}}=1.6\pm0.1$. Its low curvature is in agreement with my $^{40}\text{Ar}/^{36}\text{Ar}$ data, but it fails to connect the measured $^{38}\text{Ar}/^{36}\text{Ar}$ ratios with the solar composition. **b**, The $^{38}\text{Ar}/^{36}\text{Ar}$ versus $^{40}\text{Ar}/^{36}\text{Ar}$ diagram shows no correlation, indicating that atmospheric and mantle reservoirs have the same $^{38}\text{Ar}/^{36}\text{Ar}$ ratio. In such a three-isotope plot, two-component mixing is represented by a straight line, which can be fitted more reliably than a hyperbola.

Table 1 Neon and argon abundances ($\times 10^{-12}$ cm³ per gram at standard temperature and pressure) and isotopic ratios in MORB 2IID43

Step	^{22}Ne	$^{20}\text{Ne}/^{22}\text{Ne}$	$^{21}\text{Ne}/^{22}\text{Ne}$	^{36}Ar	$^{38}\text{Ar}/^{36}\text{Ar}$	$^{40}\text{Ar}/^{36}\text{Ar}$
3 ×	1,870	9.860 (26)	0.03049 (53)	26,400	0.18836 (33)	1,317 (56)
6 ×	71	11.785 (22)	0.05191 (40)	940	0.18836 (31)	15,700 (900)
10 ×	24	12.476 (23)	0.05890 (38)	290	0.18790 (57)	24,800 (1,100)
15 ×	5	12.446 (49)	0.05860 (66)	56	0.18784 (51)	23,800 (1,800)
40 ×	17	12.438 (26)	0.05839 (41)	240	0.18785 (58)	22,500 (1,100)
100 ×	8	12.270 (41)	0.05580 (48)	20	0.18910 (33)	18,900 (700)
400 ×	2	12.530 (170)	0.06020 (220)	30	0.18831 (57)	21,600 (1,200)
Atmospheric		9.800 (77)	0.02900 (20)		0.18800 (30)	295.5 (0.5)
Solar		13.800 (100)	0.03280 (50)		0.1724–0.1786	~0

Step values refer to the cumulative number of strokes. Numbers in parentheses indicate 1 σ error in the last digits; errors in abundances are 5% (1 σ). Values for atmospheric and solar composition from ref. 1 are for comparison.

direct evidence for such a 'polluting' mechanism might come from a correlation between maximum $^{40}\text{Ar}/^{36}\text{Ar}$ with radiogenic lead isotope ratios in MORBs¹¹.

Joachim Kunz

Institut de Physique du Globe de Paris,
4 Place Jussieu, 75252 Paris Cedex 05, France
e-mail: kunz@ipgp.jussieu.fr

1. Pepin, R. O. *Nature* **394**, 664–667 (1998).
2. Burnard, P., Graham, D. & Turner, G. *Science* **276**, 568–571 (1997).
3. Valbracht, P. J., Staudacher, T., Malahoff, A. & Allègre, C. J. *Earth Planet. Sci. Lett.* **150**, 399–411 (1997).
4. Niedermann, S., Bach, W. & Erzinger, J. *Geochim. Cosmochim. Acta* **61**, 2697–2715 (1997).
5. Moreira, M., Kunz, J. & Allègre, C. J. *Science* **279**, 1178–1181 (1998).
6. Poreda, R. & Farley, K. *Earth Planet. Sci. Lett.* **113**, 129–144 (1992).
7. Marty, B. *et al.* *Earth Planet. Sci. Lett.* **166**, 179–192 (1998).
8. Kunz, J., Staudacher, T. & Allègre, C. J. *Science* **280**, 877–880 (1998).
9. York, D. *Earth Planet. Sci. Lett.* **5**, 320–324 (1969).
10. Staudacher, T. & Allègre, C. J. *Earth Planet. Sci. Lett.* **89**, 173–183 (1988).
11. Sarda, P., Moreira, M. & Staudacher, T. *Science* **283**, 666–668 (1999).

Singing and hearing in a Tertiary bushcricket

Communication organs are poorly represented in the fossil record, so their evolution is usually reconstructed by comparison of extant species using a phylogenetic approach. We have analysed some extremely well preserved stridulatory and hearing organs of the oldest known bushcrickets from the lowermost Tertiary sediments of Denmark (55 million years old). These fossils indicate that males sang with a broadband frequency spectrum, and it is likely that both sexes could hear ultrasound. The fossil wings have lower asymmetry than extant species, indicating that bushcrickets may have evolved from a bilaterally symmetrical ancestor.

Only a few, poorly preserved male forewings from fossil bushcrickets have previously been described¹. Our specimens are from *Pseudotettigonia amoena*, which was a large bushcricket with a wing span of more than 120 mm that lived in subtropical Scandinavia during the Palaeogene². They include 11 left and right male forewings and three forelegs with the ear.

The evolution of courtship behaviour, with males producing complex songs, has led to a strong left–right asymmetry of the forewings in extant male bushcrickets³. During singing, the left wing is always held on top of the right, and the 'stridulatory file', which is a dense row of cuticular teeth on the ventral surface of the left wing, is scratched over a raised vein on the right forewing. A thin membrane called the 'mirror', which is important for sound radiation and is surrounded by a frame of strong

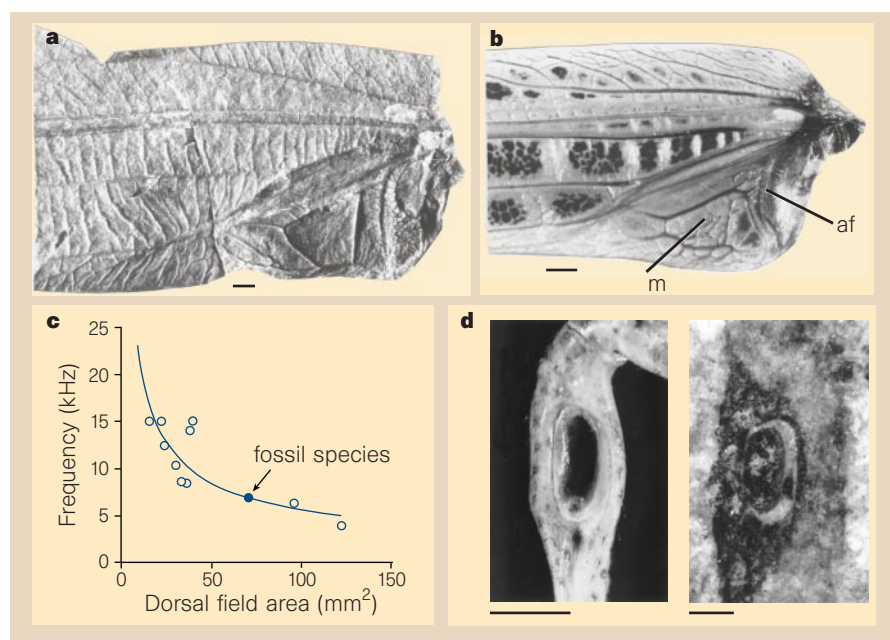


Figure 1 Comparison of the Lower Tertiary *Pseudotettigonia amoena* with a Recent Tettigoniidae. **a**, Dorsal view of *Pseudotettigonia* left forewing base (specimen HM 14M-3022) showing the stridulatory apparatus. **b**, Dorsal view of Recent *Decticus verrucivorus* left forewing base, showing the mirror (m) and the active file (af). Scale bars, 1 mm. **c**, Relation between dorsal field area and dominant frequency peak ($y = 93.886x^{-0.618}$, $R^2 = 0.79$). Data include species from extant Tettigoniidae, Phaneropteridae and Ephippigeridae^{4–6}. **d**, Tibial tympanum of fossil *Pseudotettigonia* (right, specimen HM 14M-C3272) compared with an anterior tympanum of Recent *Phaneroptera falcata* (left). Scale bars, 0.5 mm.

veins, is expressed more strongly on the right wing than on the left, and a complex of stout spines occurs exclusively on the dorsal surface of the right wing.

The fossil forewings of males show all these structures of their extant descendants (Fig. 1a,b) and are also asymmetrical. The stridulatory file is much more pronounced on the left wing, indicating that only the left file was used for sound production. The left mirror of *Pseudotettigonia* is larger than the right one, although the difference in size is less pronounced than in extant species. The spine complex occurs on the dorsal surface of both wings, in contrast to all living male bushcrickets. We interpret this condition as a primitive state in the development of wing asymmetry, and it is possible that *Pseudotettigonia* was able to fold the wings back in both left-over-right and right-over-left positions.

Comparative morphometric analysis of the dorsal fields of the fossil wings (the part of the wings that covers the body dorsally when the wings are closed) with data from Recent species^{4–6} indicates that *Pseudotettigonia* produced a broadband frequency song with a dominant frequency peak at about 7 kHz (Fig. 1c) and an ultrasonic range, which was probably less pronounced than in extant species, as *Pseudotettigonia* had a comparatively small mirror area. The size of the mirror has increased during the evolution of the stridulatory structures, probably to increase the efficiency of ultrasound radiation^{4,7}. In female bushcrickets,

most of which are silent, the mirror is absent.

The ear in the foreleg of *Pseudotettigonia* resembles the structure of modern Phaneropterinae with open tympana⁸ (Fig. 1d). Although the internal part of the hearing system is not preserved, we conclude from the modern arrangement of the different areas in the ear⁹ that the hearing range was adapted to its own song frequencies, as it is in extant species. As the fossil bushcrickets could therefore presumably hear at least low ultrasound, they should also have been able to hear the echolocation calls of bats, which first occur in the fossil record at the same geological age¹⁰.

Jes Rust, Andreas Stumpner,
Jochen Gottwald

Institut für Zoologie und Anthropologie der
Universität Göttingen, Berliner Strasse 28,
37073 Göttingen, Germany
e-mail: jrust@gwdg.de

1. Sharov, A. G. *Trudy Paleontol. Inst. Akad. Nauk SSR* **118**, 1–217 (1968). (English translation: *Phylogeny of the Orthopteroidea*; Israel Progr. Sci. Transl., Jerusalem, 1971).
2. Larsson, S. G. *Bull. Geol. Soc. Den.* **24**, 193–209 (1975).
3. Schumacher, R. *Zool. Jb. Physiol.* **82**, 45–92 (1978).
4. Keuper, A., Weidemann, S., Kalmring, K. & Kaminski, D. *Bioacoustics* **1**, 171–186 (1988).
5. Kalmring, K., Keuper, A. & Kaiser, W. in *The Tettigoniidae. Biology, Systematics and Evolution* (eds Bailey, W. J. & Rentz, D. C. F.) 191–216 (Springer, Berlin, 1990).
6. Heller, K.-G. *Bioakustik der Europäischen Laubheuschrecken* (Magraf, Weikersheim, 1988).
7. Bailey, W. J. *J. Exp. Biol.* **52**, 495–505 (1970).
8. Rentz, D. C. F. *Aust. J. Zool.* **27**, 991–1013 (1979).
9. Bangert, M. *et al.* *Hear. Res.* **115**, 27–38 (1998).
10. Stucky, R. & McKenna, M. C. in *The Fossil Record Vol. 2* (ed. Benton, M. J.) 739–771 (Chapman & Hall, London, 1993).

A user's guide to two minutes of totality

Total Eclipses: Science, Observations, Myths and Legends

by Pierre Guillermier and Serge Koutchmy
Springer/Praxis: 1999. 239 pp. £19, \$29.95

Solar Eclipse

by Thomas Crump
Constable: 1999. 256 pp. £9.99 (pbk)

Eclipse! The What, Where, When, Why & How Guide to Watching Solar and Lunar Eclipses

by Philip S. Harrington
Wiley: 1997. 280 pp. £11.99, \$16.95

The Sun in Eclipse (Practical Astronomy)

by Michael Maunder and Patrick Moore
Springer: 1997. 217 pp. \$29.95, £20

UK Solar Eclipses from Year 1

by Sheridan Williams
Clock Tower Press: 1996. 100 pp. £11.95

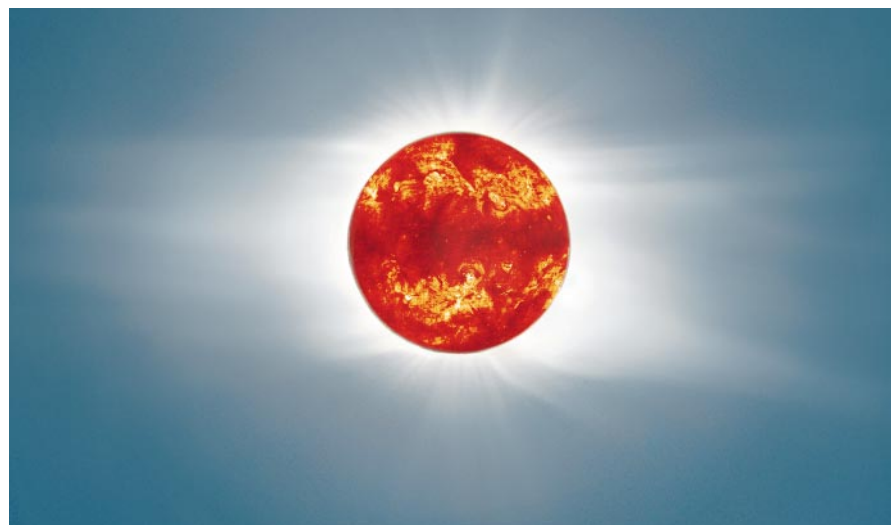
Jay M. Pasachoff

The total eclipse of 11 August that will follow a path from the Atlantic through Europe and over to India will be, for those fortunate enough to see it, one of the most wonderful experiences of their lives. For two minutes and some seconds in Europe, the blue sky will disappear, revealing the glories of the solar corona that it normally hides.

The importance of these two minutes, and the fact that one can, and should, look at the corona during this time without any filters, is unfortunately often hidden by the marketing blitz for eye-protection filters. These filters are fine in their place, but they are only for taking occasional glances at the partial phases. Indeed, putting them into eyeglass form, instead of rectangular sheets, may encourage people to use them for too long. Further, since the partial phases are not fundamentally interesting to look at, one can do almost as well by looking at the pinhole images of the crescent Sun that will be projected beneath trees and bushes.

There are several good books to choose from for background information on solar eclipses. *Total Eclipses*, a collaboration between an amateur astronomer, Pierre Guillermier, and one of the most active and enthusiastic professional eclipse scientists, Serge Koutchmy, gives a variety of information on observing eclipses for the novice and on the value of eclipses to professionals. The book, which is well illustrated, gives an authoritative survey of solar astronomy and puts solar eclipses in their scientific place.

It covers the solar survey; how to observe the Sun both inside and outside eclipses; the celestial motions that cause eclipses; historical and mythological eclipse stories; a description of what it is like to be at an



Do look now: only during totality is it safe to look at the solar corona with unprotected eyes.

eclipse; and reliable information on how to photograph solar and lunar eclipses. One of the appendices, obviously written by Koutchmy, discusses eclipse research. Another gives detailed information on this summer's eclipse. In spite of the perhaps inevitably disjointed collaboration between an amateur and a professional (a fault shared by my own co-authored *Cambridge Eclipse Photography Guide*, with Michael A. Covington; Cambridge University Press), readers can find information at an interesting and appropriate level and can be sure that their guides are knowledgeable.

Guillermier and Koutchmy discuss not merely famous eclipses in history and literature, but even try to track down an eclipse described in a Tintin story. Observers will benefit from a phase-by-phase set of recommendations on how to observe and photograph the eclipse. A variety of space observations, which are now capable of observing many aspects of the solar corona, are also discussed, including those of the current Solar and Heliospheric Observatory (SOHO). Unfortunately, the book (as of the proof stage) has not been updated since the original French edition, and so the remarkable high-resolution coronal images taken by the Transition Region and Coronal Explorer (TRACE) do not appear.

Thomas Crump, in *Solar Eclipse*, sets out to describe this eclipse, and eclipses in general, from an anthropological stance. He gleans his information from a variety of sources, necessary because he has never seen an eclipse himself. So, although many of the stories are interesting, flaws creep in, such as a comment confusing the formula for the Balmer series, which applies only to hydrogen and a few other one-electron ions, with the wavelengths of all the Fraunhofer lines.

More serious are the gaps in the single page of warning, at the beginning of the book, about viewing and photographing the eclipse. For observing information, it is better to rely on experienced observers.

Nevertheless, the stories and explanations about eclipses make the book worth reading. I hadn't known the background of horseracing's Eclipse Stakes, which is named after the racehorse Eclipse, foaled on the date of a solar eclipse and owned by King George IV when he was Prince of Wales.

One chapter is devoted to three scientific investigations traditional of eclipses: the history of 'coronium', and how it became known to consist of gases like iron at temperatures of millions of degrees, is well told. The story of the tests of Einstein's general theory of relativity at the 1919 eclipse remains of interest, even though observations using other methods, most recently with the star positions measured by the European Space Agency's Hipparcos spacecraft, give the result more accurately and obviate the need for further eclipse observations of the deflection of light. And the link between historical eclipse observations and the rotation of the Earth, perhaps leading to improved geophysical understanding of such terrestrial processes as post-glacial uplift, closes the chapter.

Philip Harrington, an experienced observer, provides a straightforward guide to eclipse watching in *Eclipse!* With only a nod to history, he moves to a clear and simple description of how to observe and photograph eclipses, also describing the darkroom and computer techniques that can be used. The heart of the book is a set of descriptions and maps of the paths from which solar and lunar eclipses can be seen between 1998 and 2017, the latter being the date of the next total

solar eclipse visible from the United States. The solar eclipse maps show only the paths of totality, and not the extent of partial phases, but are fascinating nevertheless and are all that is needed for those of us who will travel to view totality. Many of us are already plotting how to view the 23 November 2003 total eclipse that crosses land only over Antarctica.

In *The Sun in Eclipse*, Michael Maunder and Patrick Moore have written a very practical set of instructions for observing and photographing the eclipse, which fits well in Moore's "Practical Astronomy" series. The information obviously comes from knowledgeable and experienced observers. The discussion of filters and projection methods for solar observation, for example, goes on for 16 pages, and includes careful descriptions of the best filters, a discussion of pin-hole projection and a list of "filters which must never be used". Detailed hints on focusing and on the problems of camera shake are wise.

A table lists future eclipses until 2020, but the only map for 11 August is that of totality over Cornwall and Alderney, even though it is so easy to travel from England to the Continent, where weather statistics are more favourable. It is hard to understand why so many people are devoted to seeing the eclipse in England instead of crossing the channel.

Moore has a distillation of his discussion in *Patrick Moore's Guide to the 1999 Total Eclipse* (Macmillan/Boxtree), whose proceeds go towards building a planetarium in West Sussex. Maps include small versions of the standard ones by Fred Espenak of NASA; they are handy to have, but are not well reproduced, being small and on inexpensive paper.

Sheridan Williams's idiosyncratic *UK Solar Eclipses from Year 1* provides maps of all eclipses crossing England over three millennia, and shows how to observe them. Any eclipse buff should have one. *Totality: Eclipses of the Sun* (2nd edn) by Mark Littman, Fred Willcox and Fred Espenak (Oxford University Press) is late in the publication of eclipse books, but is by very experienced buffs.

The few professional teams of eclipse observers will be overwhelmed by the amateur astronomers and, even more, by the many tourists who will travel into the band of totality. Let us echo Edmond Halley's sentiment from his map predicting the path of the 1715 eclipse over England: "We wish our Astronomical Friends a Clear Sky." □

Jay M. Pasachoff is in the Department of Astronomy, Williams College, 3 Lab Campus Drive, Williamstown, Massachusetts 01267, USA, and is Chair of the Working Group on Eclipses of the solar commissions of the International Astronomical Union.

An awkward dinner guest

Mapping the Mind

by Rita Carter
Weidenfeld & Nicolson: 1998. 224 pp.
£25, \$29.95

Jeffrey Gray

In the early 1970s I had a very bright student, John Churcher, who started doing experimental psychology, then quit to study Chinese, only to come back a term later to psychology. His reason for returning was that the experience of psychology had made it impossible for him to take part in standard dinner-party conversation. His assumptions were just too different from those of everyone else. (In those days there were very few students of psychology or neuroscience.)

I knew just what John meant: it was not so much a question of having different information to bring to bear on the topic of the evening, much more to do with the conceptual framework within which it was approached. Thirty years on, despite the extensive marination of contemporary society in psychology and neuroscience, the 'Churcher gap' is still with us. So, will Rita Carter's beautifully produced and entertainingly written survey of the 'landscape of the brain' help us cross it?

Information there is aplenty, nugget after nugget of it to stimulate and tease the most jaded palate. Many of the nuggets come from recent imaging studies of the human brain in action (prettily illustrated with pictures of brains lighting up here and there in response to this and that). The attentive reader will learn to name particular regions of the brain (amygdala, hypothalamus, hippocampus...) and to attach neat labels to them (fear, sex, memory...). The more attentive reader might, however, wonder whether this kind of paired associate learning throws any real light on the way the brain actually works.

Other nuggets come from the strange vicissitudes that befall people who have had damage to discrete parts of the brain — the kind of case study that Oliver Sacks has made his own — or from various visual and other illusions taken from the treasure-house of psychophysics. Carter does an excellent job of putting together these different kinds of nugget into a unified, diverting and fast-moving narrative.

But, but... the more I read this book, the more concerned I became about the effect it will have on dinner-party conversation. It is, of course, the job of scientific journalism to make science easy to digest, but Carter makes it all too easy. Most of her story is presented as though it were

established fact, but very little of it is; and rarely is the reader given any clues for distinguishing well-supported evidence from mere speculation. Above all, one is given the impression that wonderful new machines like PET scanners have opened up startling windows on the brain/mind — lo! We now just have to look in to see what is going on there.

There is hardly any discussion of the way in which behavioural neuroscience sifts evidence, of the conceptual frameworks it uses or the problems they present, or of the difficulties of integrating theory and evidence that come from different starting points. This last issue is fudged to the point of obscurity. There are, for example, long passages that purport to be about brain function, but which are, in fact, filched from common experience and reclothed in spurious neuroscientific talk that only the worker in the field will know has no empirical support.

The conceptual confusions that permeate the book become most worrying when we reach those most likely dinner-party topics: consciousness, free will and the responsibility for one's actions. Neural messages are repeatedly sent from the 'conscious brain' to the 'unconscious brain', or in the other direction, in a way that builds shamelessly upon Cartesian dualist intuitions which are then in the final chapter piously — but without good reason given — dismissed. (The 'conscious brain', by the way, is usually said to lie in the cerebral cortex, sometimes specifically in the frontal cortex; this hypothesis, even if made intelligible, is as yet unsupported by any serious evidence. Equally unsupported is its complement, that everything below the cortex makes up the 'unconscious brain'.)

The book relentlessly blurs the critical divide between frank damage to the brain as the clear cause of behavioural derangement and variation in the activity of the brain as a correlate of behavioural variation (a correlation from which, of course, no direction of causality can be inferred). So, towards the end of the book, we are asked to take seriously the hypothesis that the massacres of Armenians by Turks, Jews by Nazis, Cambodians by Pol Pot and Tutsis by Hutus were all brought about by "spasms of overactivity in the orbito-frontal and medial prefrontal cortex".

This absurd extension of neuroscience into history and politics is what I dread most to find at the next dinner party. The Churcher gap, alas, is alive and well. □

Jeffrey Gray is at the Institute of Psychiatry and ReNeuron Ltd, De Crespigny Park, Denmark Hill, London SE5 8AF, UK.

● *Mapping the Mind* was short-listed for the 1999 Rhône-Poulenc Prize for Science Books — see page 654.

A theory that's hard to digest

The Hunting Apes: Meat Eating and the Origins of Human Behavior

by Craig Stanford

Princeton University Press: 1999. 245 pp.

\$24.95, £14.95 (pbk)

Christophe Boesch

The quest to understand what makes us human and how we became so is one of the most controversial and fascinating scientific adventures. Many scientific disciplines, including anthropology, psychology, biology and philosophy, join in this debate. The fossils of our ancestors have made it clear how closely related they were to chimpanzees. Two particularly striking differences, however, became apparent quite early in evolution: an upright walking gait and a big brain. Since most of us think that there must be more to being human than walking upright, the quest to understand human uniqueness has centred around our ancestors' development of an enlarged brain.

Craig Stanford uses some observations of wild chimpanzees to suggest a new approach to this question. Classically, anthropologists have relied on fossil bones and the artefacts associated with them. This approach gives us limited information about our ancestors' behaviour, forcing us to make inferences. Others have used knowledge of contemporary hunter-gatherer communities to make such inferences, based on the assumption that we had a similar lifestyle for most of our history. However, our ancestors were morphologically very different from modern humans. Thus, an additional source of comparison is the information we have on chimpanzees, to which our ancestors were physically similar. Stanford's attempt to add some of our growing knowledge on chimpanzees to this debate is to be welcomed.

Jane Goodall began her pioneering work on the chimpanzees in the Gombe Stream National Park, Tanzania, some 39 years ago. Her observations of the Gombe chim-

panzees using and making tools and hunting small monkeys for meat profoundly changed our view of the human-chimpanzee differences. Throughout the book, Stanford takes chimpanzee behaviour to be Gombe chimpanzee behaviour, but this produces a distorted image of what chimpanzees really are.

As Michael Tomasello and I have suggested, we should speak of different chimpanzee cultures from both biological and psychological points of view. William McGrew has shown that the tool-using habits of chimpanzee populations differ greatly; the more populations we study, the more diverse these sets of behaviour become. Similarly, many authors have shown that hunting and many other aspects of social behaviour differ strikingly between chimpanzee populations.

Chimpanzees in the Tai forest, Ivory Coast, search actively for prey, hunt in groups and coordinate their actions for most hunts, but chimpanzees at Bossou in Guinea in West Africa hunt rarely and solitarily by 'collecting' passing prey. Chimpanzees at the Mahale Mountains, Tanzania, like those at Gombe, hunt quite frequently, but rarely coordinate their actions. The diversity of chimpanzee cultures contrasts with what has been seen in other animals, but is reminiscent of humans. Whatever aspects of chimpanzee behaviour we look at, we see important differences between populations which do not reflect genetic, or sometimes ecological, differences.

The 'Man the Hunter' theory of human evolution, proposed in the early 1960s, that Stanford seeks to resuscitate, has been much criticized for ignoring women. The chimpanzees of Gombe differ from other populations, such as the Tai chimpanzees, by having females with low sociality and seemingly limited choice in their reproductive strategy. In the Tai forest, female chimpanzees are highly social, play a decisive role in many aspects of social life and make their own choice of male partners. The differences between chimpanzee cultures make us realize the flexible but important role of females in both chimpanzee and human evolution.

In addition, I don't expect that just one activity — namely hunting — could explain the evolution of humans. Humans and chim-

panzees are different from other species in many ways, and it seems reasonable to expect that other activities, such as the use and the making of tools and the use of complex social strategies, would influence our evolution. Females are prime tool users in most chimpanzee populations, and the increasing knowledge of tool use in chimpanzees reveals how complex it can be. The role in our history of females in this domain is undeniable.

What made our ancestors develop such a big brain? Stanford proposes that it was the machiavellian demands of meat-sharing: not hunting itself, and not making and using tools. However, we are left puzzled as to what about meat-sharing is so demanding that it produces the modern human mind. Large brains could mean large memories or sophisticated cognitive abilities. The skills that required such an enlargement of the brain are not stated, and we can only speculate about them. Many authors have tried to define our cognitive uniqueness, and the responses propose that it lies in the domain of abilities to understand complex causal relations about external objects, or in our ability to understand the minds of others. Meat-sharing, which in humans is often viewed as tolerated theft, is not really a convincing candidate as the driving force behind higher cognitive abilities. "What makes us human?" and "How did we become so?" remain among the most exciting questions to be studied, and many more books will have to be written to find convincing clues. □

Christophe Boesch is at the Max-Planck Institute for Evolutionary Anthropology, Inselstrasse 22, 04 103 Leipzig, Germany.

Courting success for the future

Eco-pragmatism: Making Sensible Environmental Decisions in an Uncertain World

by Daniel A. Farber

Chicago University Press: 1999. 216 pp.

\$23, £18.50

Calestous Juma

The intensity of public debate in Europe on the risks to human health and the environment posed by genetically modified crops illustrates the widening gap between public concern about environmental issues and the capacity of regulatory and judicial institutions to deal with such challenges.

There is a general crisis of confidence, and environmental concerns are becoming a prominent part of national political platforms in many countries: the public wants a fundamental reform to the way government institutions handle environmental issues. But institutional changes have been slow, and in many countries reforms have run



Share and share alike: were the demands of meat-sharing behind the evolution of the big human brain?

counter to environmental aspirations. Environmentalism is no longer a fringe activity confined to a small group of activists challenging the status quo, but a growing movement that is influencing court decisions. Indeed, this influence extends to laws and national constitutions around the world. Underlying these reforms are efforts to manage risk and uncertainty in a complex world.

The 1990s have witnessed attempts to integrate environmental considerations into institutional behaviour at national and international levels. International environmental conferences and treaties are regularly used to set the agenda for national action. But if we are to achieve this successfully, we must integrate environmental concerns into the functioning of our judicial systems.

In many cases, the judicial system has been called upon to decide between seemingly competing economic and ecological perspectives. Much of the emerging research suggests there is a false dichotomy between environment and economy. Eco-pragmatism adds to this research by seeking to establish a middle road between "tree-huggers" who support environmental conservation as a primary public-policy goal and "bean counters" who argue for the strict application of cost-benefit analysis.

Daniel Farber's is therefore a timely and well-argued contribution to the understanding of how judicial systems can maintain the balance between competing environmental and economic claims. It does more than that: it puts forward a framework for using judicial pragmatism to integrate environmental considerations into economic activities.

Farber uses the case of the Reserve Mining Co. versus the United States government to show how a pragmatic approach makes it possible to promote environmental conservation while taking into account economic considerations. In this case, the court had to decide whether to rule that a mining company should spend US\$250 million on removing a possibly carcinogenic chemical from a city's drinking water.

Arguing that "sanctity of human life is of too great value", the judge ordered the immediate shut-down of Reserve Mining. An appeal court ruled that the evidence "does not support a finding of substantial danger", and the discharges posed only a "possible medical danger". The matter then turned to how to identify a feasible land-disposal method.

The book shows how the courts translated the environmental standards embedded in American culture into decisions that promoted conservation while applying cost-benefit analysis. It is a compelling account of how the judicial system dealt with a complex case under conditions of uncertainty and growing scientific knowledge. Generalizing from this account, Farber proposes an approach of cost-benefit analysis and feasibility

analysis designed to reflect environmental considerations.

Farber admits that "there are limits to how far Americans are willing to go to achieve environmental goals". This lack of willingness itself requires explanation because of its implications for relations between the United States and the rest of the world. Many argue for approaches governed by caution. Given the current sweep of globalization, a dialogue between different value systems is essential. One of the key interna-

tional goals today is to reconcile pragmatism with the precautionary approaches that dominate European cultures.

A pragmatic approach should lead to an ecological renaissance and contribute to our understanding of how our existing cultures (and the institutional arrangements that come with them) influence our approach to environmental issues. □

Calestous Juma is at the Center for International Development, Harvard University, Cambridge, Massachusetts 02138, USA.

Hoffman's Erdős is number one

Paul Hoffman's book *The Man Who Loved Only Numbers* was announced as this year's winner of the Rhône-Poulenc Prize for Science Books at a dinner at London's Science Museum last week. Hoffman was one of six finalists:

The Man Who Loved Only Numbers: The Story of Paul Erdős and the Search for Mathematical Truth

by Paul Hoffman

Fourth Estate, £7.99, \$12.95 (pbk)

"... Hoffman is very good at discussing the sort of problems that intrigued Erdős, which makes *The Man Who Loved Only Numbers* an excellent introductory book for anyone interested in number theory. As a subject, however, Erdős is difficult. He led a very determinedly undramatic life and Hoffman is forced to fill up many pages with unrelated detail and stories about mathematicians. This upsets the flow of the biography, but frequently adds to the book's interest... *The Man Who Loved Only Numbers* is a fascinating, affectionate biography, but the title is untrue: Erdős loved many things besides numbers, including children (whom he called epsilons, after the mathematical symbol for a very small quantity) and especially his mother, who accompanied him on many of his early travels." Alexander Masters, *Nature* 394, 535-536 (1998).

A Beautiful Mind

by Sylvia Nasar

Faber, £17.99

This book will be reviewed in *Nature* in July.

How the Mind Works

by Steven Pinker

Norton/Penguin, \$17, £9.99 (pbk)

"... Pinker has a gift for exposition, for witty analogies, for apposite quotation from his vast knowledge of culture, high and low. And he knows when to skip the details of difficult technicalities. He uses all these skills in his latest book. It is erudite, light in touch, but long... The most controversial aspect of Pinker's package is his defence of evolutionary psychology... Pinker's evolutionary psychology is useful heuristic for generating hypotheses, but not a refutable theory." P. N. Johnson-Laird, *Nature* 389, 557-558 (1997).

Consilience: The Unity of Knowledge

by E. O. Wilson

Knopf, \$26

"... Its first thesis is that the humanities, sociology, religion, ethics, art appreciation and almost everything else outside the current remit of science that has not found its analytical roots in evolutionary biology will soon do so. Its second thesis is that, once we understand how we came to be, we shall be constrained in determining where to go. Wilson makes a formidable argument for consilience — unifying on the basis of a common theoretical framework — of those near-disciplines that have no rational structure... Wilson is a kind man, but he has no problems producing withering denunciations where they are called for. His description of our ecological plight, and his heartfelt plea for understanding and urgent action, are written with such authority, clarity and passion that I have read nothing to touch them." Paul H. Harvey, *Nature* 392, 451-452 (1998).

One Renegade Cell: The Quest for the Origins of Cancer

by Robert Weinberg

Weidenfeld & Nicolson, £12.99

Robert Weinberg's brisk trot through the scientific efforts to understand cancer joins the recent spate of books in which experts write a history of their own field. It is amazing science, but the tired metaphors and unwieldy gene names will not make it easy reading for lay-readers even if it is all terribly relevant to them — as Weinberg points out, about 40 per cent of the population of the United States will develop cancer at some stage in their lives. Nor does the ultimately unsettling subject matter recommend the book as one to read in bed. — Harriet Coles

Mapping the Mind

by Rita Carter

Weidenfeld & Nicolson, £25

See page 652 for a review.

Topography of contextual modulations mediated by short-range interactions in primary visual cortex

Aniruddha Das & Charles D. Gilbert

The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

Neurons in primary visual cortex (V1) respond differently to a simple visual element presented in isolation from when it is embedded within a complex image. This difference, a specific modulation by surrounding elements in the image, is mediated by short- and long-range connections within V1 and by feedback from other areas. Here we study the role of short-range connections in this process, and relate it to the layout of local inhomogeneities in the cortical maps of orientation and space. By measuring correlation between neuron pairs located in optically imaged maps of V1 orientation columns we show that the strength of local connections between cells is a graded function of lateral separation across cortex, largely radially symmetrical and relatively independent of orientation preferences. We then show the contextual influence of flanking visual elements on neuronal responses varies systematically with a neuron's position within the cortical orientation map. The strength of this contextual influence on a neuron can be predicted from a model of local connections based on simple overlap with particular features of the orientation map. This indicates that local intracortical circuitry could endow neurons with a graded specialization for processing angular visual features such as corners and T junctions, and this specialization could have its own functional cortical map, linked with the orientation map.

The response of a neuron in primary visual cortex (V1) to an element of a complex image is influenced significantly by the context surrounding that element¹. The specific layout of the surrounding features can facilitate or suppress the neuronal response and alter its temporal structure markedly, compared to the response to the same element presented in isolation. It is assumed that this modulation is involved in the analysis of form. Investigation of the physiological substrates of these context-dependent interactions has focused largely on long-range horizontal connections in V1^{2–4} that are believed to be involved in perceptual grouping and contour saliency^{5–7}. Less attention has been paid to the role of local connections between neighbouring cells.

The effective local intracortical connections of a neuron—and the consequent local visual processing—are likely to depend strongly on the location of the neuron in the cortical maps of orientation and space. Maps of orientation columns over the surface of V1 show varying degrees of local homogeneity. Cortical columns where orientation preference changes smoothly or remains essentially constant over distances of 500 μm or more are interspersed with regions containing orientation singularities—reversals, sharp jumps (fractures) and point singularities^{8–11} where the orientation changes by up to 90° in less than 150 μm . The map of retinotopy on V1 may show similar inhomogeneities that correlate closely with inhomogeneities in the map of orientation¹². Neurons within a cortical column of uniform orientation have largely overlapping receptive fields, whereas neurons separated by as little as 150 μm across an orientation fracture (which have mutually orthogonal orientation preferences) have largely non-overlapping receptive fields. However, anatomical studies indicate that the dendritic and local axonal arborizations of pyramidal neurons are largely circularly symmetrical, and extend laterally out by 300–500 μm independently of the neuron's proximity to functional boundaries such as orientation singularities¹³ or other functional features within V1 (refs 14–16). Neurons close to orientation singularities are thus likely to be anatomically connected to neurons preferring a

range of orientations, including orthogonal ones, in neighbouring cortical columns¹⁷. By contrast, neurons distant from orientation singularities, lying within a region of uniform orientation, are linked anatomically with a more homogeneous population of cells having similar orientation preferences and receptive fields.

Here we explore whether these patterns of local anatomical connectivity translate to similar patterns of effective physiological connection strength, and the role that such local connections have in the visual processing of complex images. We address two questions: first, how much does the effectiveness of local physiological connections depend on cortical distance, compared to the relative preferred orientations of two cells? Second, does the pattern of local connections predict the nature of neuronal responses to complex visual stimuli, and, in particular, the influence of local visual context on a neuron's response to its preferred stimuli? For this second question we restricted our study to visual contexts provided by line elements orthogonal to the preferred orientation of the recorded neuron^{18–20}, to probe selectively the local—as opposed to long-range—cortical connectivity. The modulation of a neuron's activity by iso-oriented flanking stimuli is believed to involve long-range horizontal connections that selectively interconnect cortical columns of similar orientation preference in V1 (refs 21–24). The influence of perpendicular flanking bars is thus likely to be the least dependent on long-range circuitry¹⁷. All experiments were done in cat V1 using optical imaging of intrinsic signals (that is, changes in cortical reflectance linked to metabolic changes resulting from local neural activity^{25,26}) and electrophysiological recordings from neurons in upper layers (up to 400 μm deep).

Distribution of local connection strengths

We measured the strengths of connections between pairs of neurons as a function of cortical separation and relative orientation difference, using the measure of cross-correlation strength. Earlier experimental results and theoretical considerations indicate that the height and area of the correlation peak for a pair of neurons,

normalized by the correlation baseline, gives an estimate of the 'effective connection strength' between the neurons^{23,27–32}. Recording sites were chosen after obtaining an optical map of V1 orientation columns and the recorded neurons were driven with optimally oriented line stimuli.

Using cross-correlation, we found strong effective physiological connections between pairs of neurons of all relative orientation preferences and at lateral separations of up to about 800 μm . This was true even for pairs of neurons that had mutually orthogonal preferred orientations (Fig. 1a), with sharp orientation tuning, non-overlapping receptive fields and no common responses to single bars at any orientation (Fig. 1a). When we held one electrode fixed and moved the other to different sites, at roughly the same distance but with very different preferred orientations, we found correlation peaks that were similar in height and area (Fig. 1b), indicating that the effective physiological connection strengths between pairs of neurons at comparable lateral separations is largely independent of orientation preferences.

We controlled for the null hypothesis that only neuron pairs

with similar orientation preferences are significantly cross-correlated and that the correlation obtained for orthogonally oriented neurons was an artefact of inadequate spike isolation. Such an artefact could arise, when recording from orthogonal orientation columns, through contamination by a small fraction of spikes from cells having the same, rather than the orthogonal, orientation preference. This fraction could be small enough that orientation tunings for the two sites appear orthogonal, but the contaminating subset of spikes could be the ones actually generating a significant cross-correlogram and be responsible for the apparently strong correlations between orthogonally oriented neurons. Figure 1c shows the results of a control experiment to test for this possibility, with a pair of sites at a separation of 280 μm and relative orientation difference of 80°. One of the recording sites (red) was close to an orientation singularity, where the chance of contamination from units of different orientation preferences is high. If the correlogram obtained by optimally stimulating both of the nominally orthogonal channels (Fig. 1c, upper correlogram, not normalized by the baseline) were due to such contamination, then stimulating the

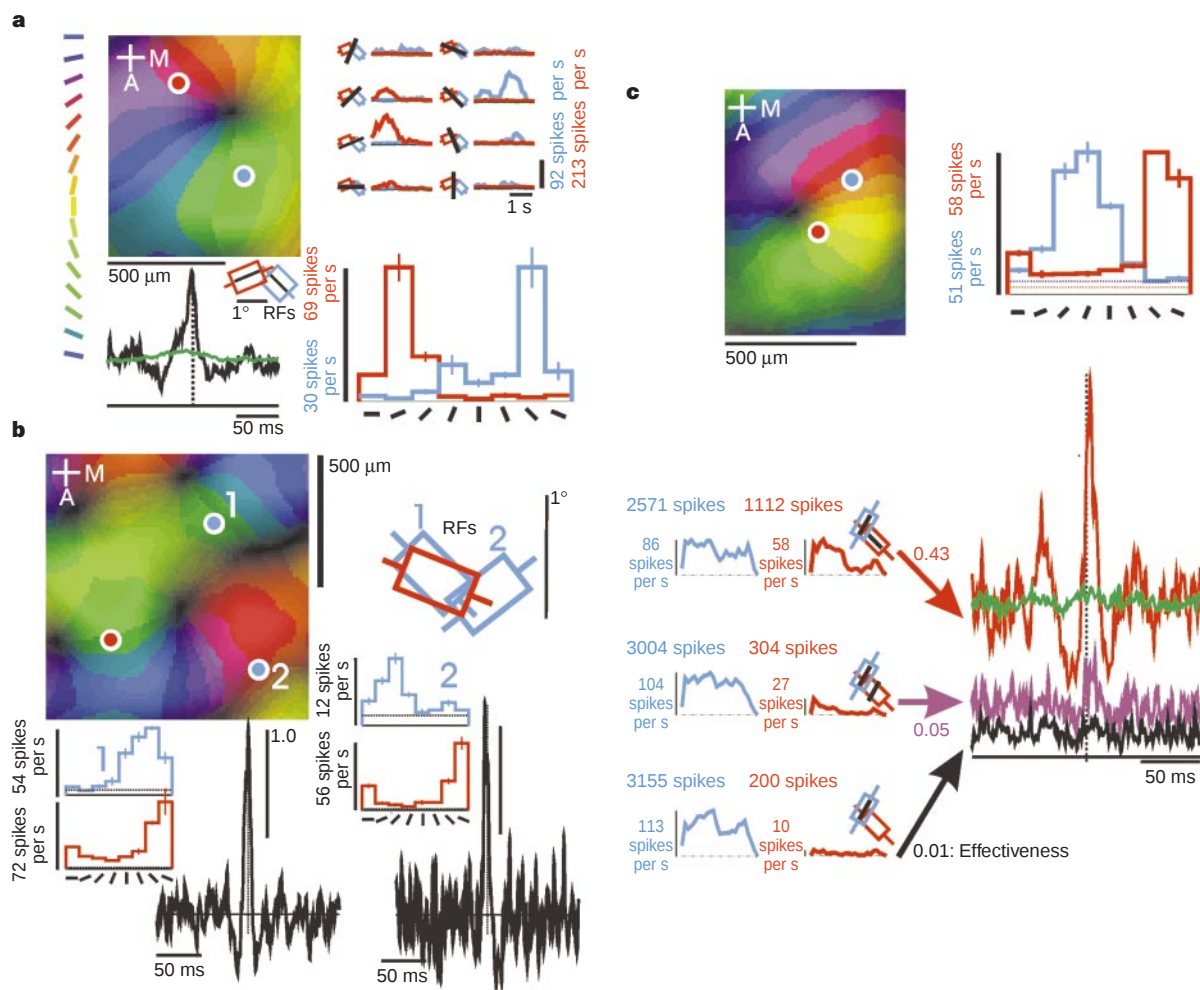


Figure 1 V1 neurons have strong local interconnections, independent of relative orientation difference. **a**, Results for a neuron pair with a 70° orientation-preference difference. Top left: optical image of orientation map with electrode recording sites. The corresponding receptive field (RF) outlines, PSTHs to single oriented bars (upper right) and orientation tuning curves (lower right) are colour-coded by recording site (blue and red). An independent colour code is used for the optical map (key to left of optical image, see Methods; black line to left of each PSTH, or below orientation tuning curve, indicates stimulus orientation). Lower left: cross-correlogram obtained with two-bar stimulus optimally driving both neurons simultaneously (black: raw correlogram; green: shuffled correlogram baseline; normalized peak height, 1.46). **b**, Comparing correlograms for neuron

pairs with one common reference (the 'red' site). The 'blue' electrode was positioned in a similarly or orthogonally oriented column (sites 1 and 2, respectively). The corresponding receptive fields, orientation tuning curves and normalized, shuffle-corrected cross-correlograms are shown. **c**, Top: PSTHs and raw cross-correlogram (green: shuffled baseline) obtained by stimulating both 'red' and 'blue' receptive fields with optimally positioned and oriented bars. Middle: the same, but with stimulus bars that were both oriented for the 'blue' site. Bottom: the same, but using a single stimulus bar optimal for the receptive field 'blue' alone. The 'effectiveness'²⁸ (total number of spikes in raw peak/number of spikes in 'blue' channel) is shown for each case.

two channels by stimuli oriented for the reference channel (middle correlogram) or stimulating the reference receptive field alone (lower correlogram) should have preferentially driven the 'contaminating' spikes. We would then expect raw correlogram peaks of similar sizes. This was not the case, indicating that the cross-correlations between orthogonally oriented neurons are not due

to such contamination. Note that the average spike rate and total number of spikes in the reference channel remained essentially the same for all three configurations of stimuli.

Our pooled results (from 85 neuron pairs) indicate that neighbouring upper-layer neurons in V1 have strong physiological connections with each other, largely independent of relative orien-

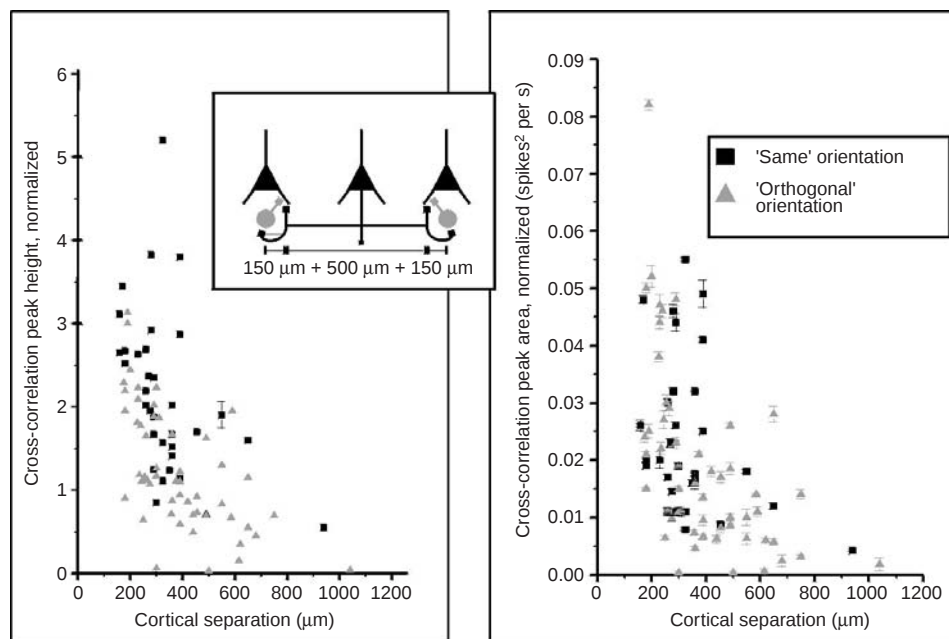


Figure 2 Normalized cross-correlation peak height (left) and peak area (right) for receptive field pairs as a function of inter-site separation on cortex. The full population is divided into 'similarly oriented' or 'orthogonally oriented' pairs. (Error bars, $\pm$ s.d. Only one (average) error bar is shown for peaks, to avoid clutter.) Inset: a plausible circuit of local interconnection between neurons separated by up to

800 μ m: a source of common input with axonal arborization extending out to 500 μ m in diameter could be connected to target neurons each with a dendritic arborization of 150 μ m radius, both directly (excitatory connection) and through an inhibitory interneuron.

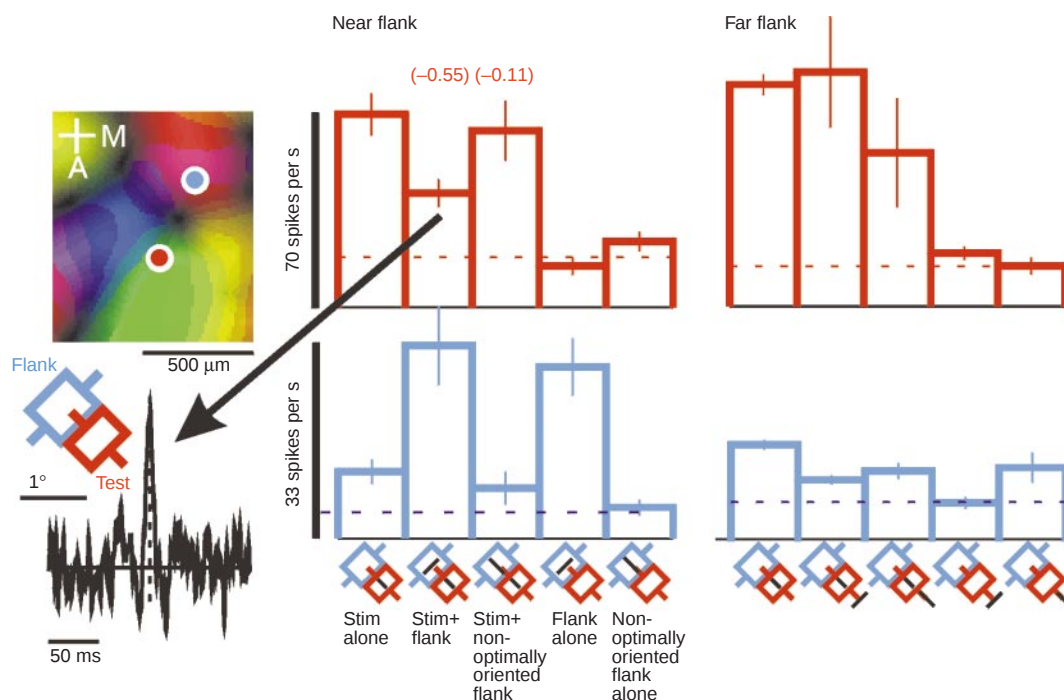


Figure 3 Specific suppressive influence of perpendicular flank stimuli. Top left: The test site (135° orientation) is labelled red; blue marks a recording site monitoring activation of the neighbouring cortical region of 45° orientation, selected as the flank. Right: Histograms of responses in the test and flank neurons. Black lines in the key below the flank response histogram indicate the

stimulus configuration for each pair of test and flank responses; numbers in parenthesis quantify test suppression (see Methods; error bar, $\pm$ s.e.m.; $n = 11$). Lower left: Cross-correlogram obtained from the spike trains stimulated by the 'Near flank', 'Stim + flank' configuration. (Correlogram shuffle-corrected and normalized: peak, 0.79; area, 0.006.)

tation preferences. The connection strength is a graded function of cortical distance, at a spatial scale consistent with anatomical connectivity. For relatively short cell separations of up to about 800 μm , the cross-correlation peak sizes, and particularly peak areas, were independent of relative orientation difference over the population (Fig. 2) and were essentially a function of inter-site separation. All the cross-correlation peaks measured had a significant value at $\tau = 0$, implying that a large portion of the synaptic input captured by the correlation process is of the common input type. A lateral separation of 800 μm , though large, is consistent with the known anatomy of local axonal connections. Pyramidal neurons in upper layers of cat V1 have local axonal arborizations up to 500 μm in diameter and dendritic arborizations up to 300 μm in diameter. Thus pyramidal neurons can provide a common input to upper layer neurons separated by up to 800 μm (inset, Fig. 2). Note that the connections providing common inputs are likely to be directly excitatory as well as having an inhibitory component through local inhibitory interneurons^{33–35} (inset, Fig. 2). The range of correlation peak heights for ‘same orientation’ neuron pairs appeared to be slightly higher than the range for ‘orthogonal’ neuron pairs, presumably reflecting narrower correlogram peaks (Fig. 2), that is, a tighter temporal correlation. The effective connection strengths between neuron pairs are, however, more reliably related to normalized correlogram peak areas, which showed little such bias overall in the population, rather than to peak heights^{27,31}.

Local circuitry and contextual modulation

The above analysis implies that the local circuits for visual processing in V1 include strong lateral connections with nearby orientation columns. Over a spatial scale of 500–800 μm , these connections are largely independent of the orientation preferences of the columns. Owing to inhomogeneities in the maps of orientation and space, which repeat at roughly the same scale of 500–800 μm , the relative orientations and receptive field positions of neurons that make up a local neighbourhood vary considerably with their positions on the cortex. We investigated the consequences, for neuronal processing of complex stimuli, of this pattern of local connectivity and its systematic variation across the cortex. As a probe of contextual modulation we used flanking perpendicular lines, lying immediately outside the classical receptive field, which influence the responses of neurons to optimal stimuli within their receptive fields. After obtaining the optical map of orientation columns we chose a ‘test’ recording site lying in a cortical column of a given orientation preference. This site was situated at a specific distance from a neighbouring cortical region of orthogonal orientation preference (the chosen ‘flank’ cortical region to be driven by flanking perpendicular stimuli). Using electrode recordings, we mapped the distribution of receptive fields represented in the local area of cortex. Short line segments of optimal size, orientation and position were then used to drive, selectively, neurons in the recording site and in the flanking region of cortex.

The primary influence of flanking perpendicular visual stimuli was to suppress the responses of neurons to their optimal stimuli, particularly when the contrast of the flank stimulus bar was more than twice that of the bar within the receptive field (Fig. 3). After optically imaging the orientation map, we positioned two electrodes in two neighbouring, orthogonally oriented cortical columns. One of the recording sites was then selected as the test site and its corresponding optimal stimulus as the test stimulus. The other neuron’s receptive field (‘flank receptive field’) was used to position the flank stimulus (Fig. 3). The flank neuron’s responses were monitored as a measure of the activity of neurons in the corresponding cortical region (the flank cortical region for this particular configuration) for different combinations of test and flank stimuli. A bar stimulus optimally oriented and positioned to drive the flank neuron led to a 55% reduction in the response of the test neuron to

its own optimal stimulus (Fig. 3, ‘Stim + flank’). Only stimuli that specifically drove the neighbouring flanking region of cortex led to such a reduction in the signal strength in the test neuron. A bar at the same position in visual space, inside the flank receptive field but no longer optimally oriented to drive flank neurons (Fig. 3,

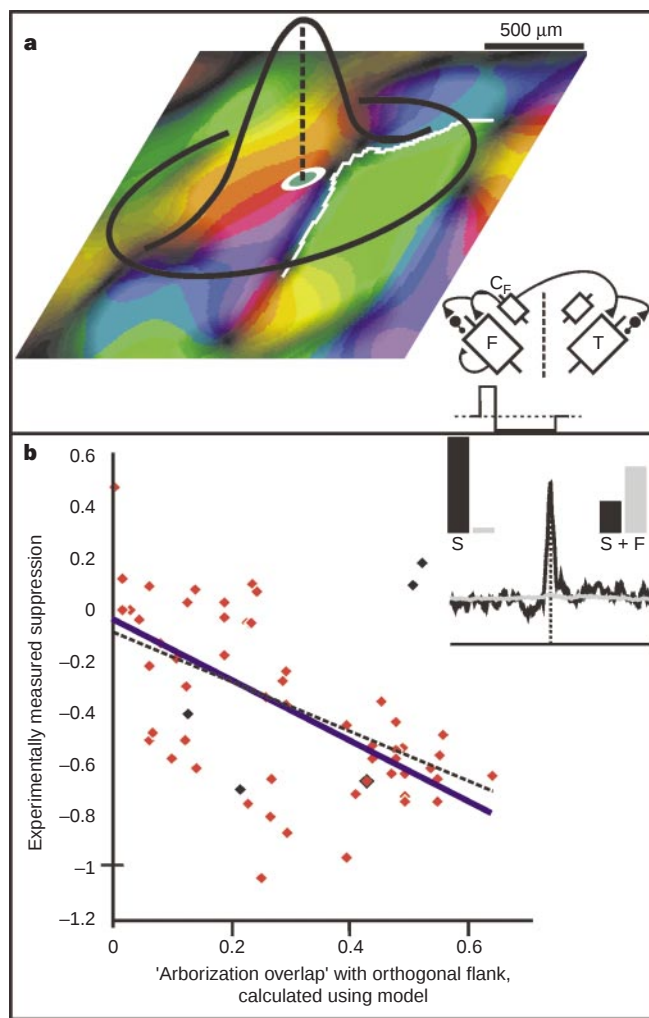


Figure 4 Comparing suppression strength with arborization overlap calculated using geometric model. **a**, Quantitative model of arborization overlap (see Methods) illustrated for a specific recording site within a 45° cortical column but close to a 135° column. The white outline marks the flank cortical region selected by the model as ‘orthogonal’ in orientation to the recording site. Inset: model to simulate suppression by flank stimuli, mediated through common input (Fig. 3; see Methods). T and F represent simultaneously recorded test and flank neurons, in orthogonal cortical columns separated by a singularity (dashed line). C_F represents other neurons in the flank column, providing direct excitatory input (triangles) and di-synaptic inhibitory input (circles) to both F and T. Units in the test column provide similar, mirror-image connections (not shown). The profile of each common-input excitatory, followed by inhibitory, modulation is indicated below the circuit. **b**, Experimentally measured suppression plotted as a function of the corresponding model, arborization overlap. (The four points in black mark the far/near pairs where ‘near’ suppression was weaker than ‘far’, Fig. 5c. Blue line: regression fit, $n = 56$, $R^2 = 0.39$, through all points excluding the four in black; black dotted line: $n = 60$, $R^2 = 0.25$, through all points.) Inset: results of simulated suppression by flank. S: simulated response histograms when only test column was activated; S + F: with simulated activity in flank neuron, the test response is suppressed. The normalized cross-correlogram of the test and flank spike trains simulated in the S + F condition is also shown. (Simulated test response histogram shown in black, flank in grey). See Supplementary Information for details.

‘Stim + non-optimally oriented flank’), had no effect on the signal. Neither did we see any suppression when moving the modulatory bar to a mirror-image position in visual space, at the same distance from the receptive field of the test neuron but far from the flank receptive field (Fig. 3, ‘Far flank’). Note that the responses of the test and flank neurons were still significantly cross-correlated—implying a strong physiological connection between them—even when stimulating the flank resulted in a suppression of the test response (Fig. 3, lower left; also see Fig. 4, insets, and Supplementary Information). The common input that presumably underlies the cross-correlation is likely to provide excitatory input through direct synaptic connections on the target neurons, whereas inhibitory input would come through an inhibitory interneuron and hence would be delayed with respect to the excitation³⁵ (inset, Fig. 2; numerical simulation in Fig. 4a, inset). This would result in an excitatory increase in common firing probability preceding any inhibitory decrease, leading to a correlation of the common input-induced spikes in the test and flank neurons, even though the net signal decrease due to inhibition could be larger than any net excitatory increase (see Fig. 4, insets, and Supplementary Information for a quantitative model of this effect).

The suppression induced by a flanking perpendicular stimulus became stronger as the test recording site came closer to the cortical region activated by the flank stimulus. This is illustrated in Fig. 5a where we compare the suppression induced by a common flanking stimulus in two separate, simultaneously recorded test neurons with overlapping receptive fields but with recording sites at different cortical distances from the flanking orthogonal cortical column. Figure 5b shows a configuration complementary to that in Fig. 5a. Here we compare the suppression induced in the same neuron by two different flank stimuli equidistant from the test receptive field in visual space but corresponding to cortical columns at different anatomical distances from the recording site. Once again, the flanking stimulus that drove cells in the cortical column anatomically closer to the test site induced a significantly stronger suppression. By plotting the two-dimensional receptive field profile for the test neuron using an automated process (Fig. 5b; see Methods) we confirmed that the two flanking stimuli were symmetrically placed outside the classical receptive field of the neuron. This controls for

the possibility that the stronger suppression caused by the ‘near’ flank could result from a trivial asymmetry in placement around the test receptive field, or to an asymmetry in the test receptive field profile leading to unequal engagement with the two flanks. In Fig. 5c we show that the pattern of flank suppression illustrated in Fig. 5a and b is mirrored over the entire population of near/far pairs recorded (31 pairs), whether with pairs of test sites that were near or far from a cortical flank column (Fig. 5a), or with flank cortical columns that were near or far from a common test site (Fig. 5b). For each pair in the population (with two exceptions: Fig. 5c) the near configuration produced a stronger suppression—by a factor of about 3 on average—than the corresponding far configuration.

We used a quantitative model of the overlap of a neuron’s arborization with nearby cortical columns to show that the measured strength of flank suppression was proportional to the expected overlap with the corresponding cortical flank (Fig. 4). By using vascular landmarks, each test site was located on the corresponding optical image of the orientation map. The basin of local interactions feeding into the neuron was modelled as a circularly symmetrical gaussian centred on the recording site and extending roughly 800 μm (see Methods for details). Figure 4a shows the model for a particular recording site. The strength of interaction with a given orthogonal cortical column was modelled as the weighted overlap of the gaussian with the area covered by the cortical column in the optical map. Larger values for this overlap represented a combination of greater proximity of the given recording site to the cortical flank in question and a larger cortical area of orthogonal orientation in the flank. Figure 4b shows the pooled results for 60 test sites, with the experimentally measured strength of flank suppression plotted as a function of the modelled arborization overlap. The measured suppression is linearly proportional to the calculated overlap ($n = 60$ recording sites in 11 animals; regression $R^2 = 0.39$ when excluding the two pairs of points in which ‘near’ suppression is weaker than ‘far’ (Fig. 5c); $R^2 = 0.25$ when including these two pairs of points). This result indicates that even a simple geometric model of arborization overlap gives a plausible measure of the strength of local interactions that could underlie the suppression of neuronal activity by nearby perpendicular flanks.

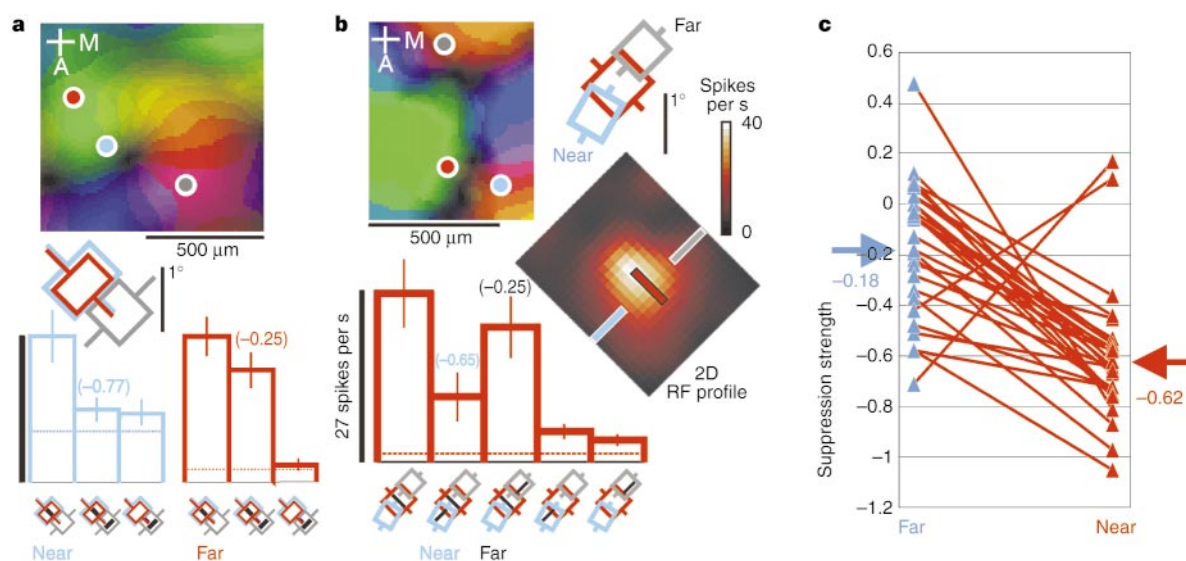


Figure 5 Flanks originating from nearer cortical loci lead to stronger suppression. **a**, Comparing simultaneously recorded test neurons, near (blue) and far (red) from the flank region (grey recording site and receptive field). The response histograms, stimulus configuration key and corresponding suppression are shown as in Fig. 3. **b**, Comparing the influence of a near (blue) and far (grey) flank,

respectively. The test (red) and the two flank stimuli are shown overlaid, to scale, on the 2D profile of the test receptive field (see Methods). **c**, Comparing suppression strength for all far/near pairs (31 pairs, 12 cats; each pair connected by a line) quantified as in Fig. 3. Arrowheads show the mean value of suppression for each condition.

Discussion

Our results have a number of possible consequences for our understanding of both the general principles of cortical circuitry and local processing, and the particular mechanisms and architecture underlying the processing of complex visual stimuli.

Our first finding is that the strength of local cortical interactions is a graded function of cortical distance with a spatial scale matching the periodicity of cortical maps and independent of relative orientation. This provides a basis for a general principle that could underlie cortical processing. The effective local functional environment and local circuits affecting a neuron in V1 would depend critically on its location within the interrelated maps of position and orientation. The map of retinotopy over the surface of V1 is locally inhomogeneous, in register with inhomogeneities in the map of orientation¹². In particular, neurons in cortical columns of mutually orthogonal orientation, lying on either side of an orientation singularity, represent non-overlapping (albeit abutting) regions of visual space. In cortical regions where orientation changes are relatively gradual, by contrast, the representation of space changes slowly. The current finding supports our proposal that neurons situated near orientation singularities could be engaged in local computations of visual features that are locally maximally dissimilar, through interactions with neighbouring orientation columns, and neurons in cortical regions of smoothly changing orientation could be engaged in computations of features that are maximally similar¹². This could be a general motif in local cortical processing, resulting from the close match between the spatial scales of local interactions and the periodicity of the cortical map. Through local circuitry, an anatomically homogenous population of cortical neurons could thus achieve the widest possible dynamic range of any given response property by being distributed over a range of positions within—and having a corresponding range of overlaps with—a full set of cortical functional columns^{36–38}. A functional ‘module’ on V1 would therefore comprise a cortical region containing a full range of such local neuronal circuits. These modules would have the same spacing over the cortical surface as the singularities in the map of orientation, with each module containing a full representation of all orientations covering a local region of visual space. These proposed modules would thus exhibit the same spacing, coverage and magnification as the ‘ice cube’ modular architecture proposed for V1 (ref. 38). Note that the cortical calculations proposed for these modules involve local neural circuitry that operates within the ambit of neighbouring cortical columns. This is to be distinguished from long-range horizontal connections over larger cortical distances that connect columns of similar orientation preference across different local modules. Visual interactions mediated by local connections are thus likely to operate at a smaller spatial scale than the processing that has been attributed to the long-range horizontal connections.

Our findings reveal a new set of response properties in V1 generated by local cortical interactions: a set of graded specializations for processing locally complex visual features. Increasing evidence indicates that cells in V1 are context dependent and that their responses may be specific for more complex stimulus configurations than a single oriented line segment¹. The pattern of perpendicular flank modulation indicates that neurons might be functionally specialized for processing corners and T-junctions of particular handedness and relative layout. Different neurons in the same orientation column—which otherwise have similar orientation preferences and overlapping receptive fields—would show different degrees of this specialization; this could therefore be another response property with which to classify receptive fields, along with orientation, ocular dominance, disparity and so on. The purely suppressive modulation by perpendicular flanks seen here could signal the presence of corners or T-junctions by terminating the facilitatory modulation that propagates along cortical representations of smooth contours². It could, alternatively, result from

this study’s being carried out in anesthetized animals, exposing a state of the cortical circuit unmodulated by attention or task demands^{17,39–41}. Note that our findings deal with the role of local interactions in mediating the contextual influence that flanking lines exert on neuronal responses to simple stimuli. Local intracortical interactions have also been postulated to play a part in the shaping of the classical response properties of neurons. It has been suggested that the sharpness and dynamics of orientation or direction tuning of neurons in the upper layers of V1 could be determined by lateral intracortical connections operating over the same lateral scales as seen here⁴². Local inactivation in V1 leads to increases in the width of orientation tuning, or changes in direction selectivity, in an orientation-dependent manner supportive of a role for inhibition in maintaining such receptive field response properties in neurons in the upper layers of V1 (refs 43, 44). The above studies deal with the forming of a neuronal response to a single line stimulus rather than the modulation of responses to a preferred stimulus through the presence of nearby flanking stimuli, and thus address an issue of distinctly different consequences for visual processing.

Like orientation preference, selectivity for more complex forms could have a systematic functional architecture. For neurons in the same orientation column, the degree of selectivity for processing corners is tied to the location of a neuron within the column, in a manner that can be predicted using a simple model of the overlap of the neuron’s arborization with local features in the cortical map. This indicates that the parameter of corner processing could also form a functional map over the cortical surface, similar to and closely linked with the familiar maps of orientation and space. The evidence for this particular functional specialization raises the question of how diverse a group of distinct contextual interactions the primary visual cortex could encode. Even a generous estimate of the number of neurons required to represent a complete set of simple visual elements of orientation, disparity and ocular dominance in a local region of visual space leaves an apparent 100- to 1000-fold redundancy in the number of cortical neurons engaged in such a representation. This apparent redundancy is progressively reduced when we consider cortical responses to more complex stimuli; adjacent neurons, although displaying similar orientation preferences, have widely divergent responses to stimuli such as checker-board patterns⁴⁵. The particular functional specialization reported here could prove to be just one of a much larger group of functional specializations mediated through local circuits and systematically organized over the surface of V1. □

Methods

Adult cats were anaesthetized, paralysed and prepared for electrical and optical recording from V1. Eyes were fixed relative to the stereotaxic apparatus by using a ring glued to the sclera.

Optical imaging. The orientation map on V1 was visualized by optically imaging the intrinsic cortical signal, as described¹². The computed orientation vector at each point is colour coded for orientation preference at 11.25° intervals starting with blue (bin 0° to 11.25°; see key to left of optical image, Fig. 1a). The brightness of each colour indicates the strength of the computed orientation vector, with dark points or lines indicating singularities or fractures.

Electrophysiology. All electrical recordings were done using pairs of independently positioned electrodes (etched tungsten in varnish), to have a reference receptive field to monitor eye position. Receptive fields were single unit, isolated by setting the level of the spike discriminator at 3× to 5× the baseline and estimating by eye the constancy of the shape of the selected waveform (amplifier: AM Systems; spike discriminator: Tucker-Davis SD1), and were obtained within the first 300 μm—typically, within the first 150 μm—of the cortical surface. Receptive field boundaries were plotted by hand as minimum response fields.

2D receptive field profile. Response histogram obtained while a single test bar was flashed in pseudo-random sequence at the vertices of an $M \times N$ (typically 6 × 5) grid centred over the hand-plotted receptive field profile.

Peristimulus time histograms and orientation tuning curves. Oriented bar stimuli (size range: $0.25^\circ \times 0.1^\circ$ to $1^\circ \times 0.25^\circ$; luminance range 7–15 cd m^{-2} against a background of 2.5 cd m^{-2}) were drifted at 1° s^{-1} and averaged over 5–15 sweeps. Orientation tuning curves were obtained by averaging over the corresponding peristimulus time histograms (PSTHs).

Cross-correlation. Calculated from spike trains of 2,000–10,000 spikes per neuron, 0.1 ms resolution, induced by drifting bars; using regions of PSTHs that were flat (within a factor of 2) for >1 s. Correlogram window, ± 100 ms, bin width 0.5 ms. Correlograms were corrected by the averaged shift predictor³¹ (also known as shuffled correlogram or baseline): $X\text{-Corr}_{\text{SHIFTED,AVG}} = \Sigma(X\text{-Corr}_{\text{SHIFTED,I}})/(N-1)$ (where $X\text{-Corr}_{\text{SHIFTED,I}}$ is the cross-correlogram obtained by shifting by an integral number I of sweeps), smoothed using a 5-bin kernel: [0.05, 0.25, 0.4, 0.25, 0.05] and normalized by the average baseline (over the correlogram window). The height and integrated area of the normalized cross-correlogram peak was used as our measure of the ‘effective connection strength’ linking the neuron pair under study^{27,28,31}. Error estimates: for the peaks, standard deviation was equal to the standard deviation in the flanks of the normalized correlogram; for the area, the variance was estimated as the variance of the normalized correlogram, integrated over the width of the peak.

Population distribution of cross-correlation strength. The full population of 85 neuron pairs was divided into two groups: ‘similar orientation pairs’ (relative orientation difference $<30^\circ$; 33 pairs) and ‘orthogonal orientation pairs’ (relative orientation difference $>60^\circ$; 52 pairs).

Quantifying suppression by flank stimulus. ((Response to Stim + flank) – (Response to Stim alone))/(Response to Stim alone), after correcting for spontaneous activity.

Simulation of suppression by flank. Test and flank neurons, as well as ‘common units’ in test and flank cortical columns, were simulated as independent Poisson spike generators with variable rate. Spikes in either ‘common’ unit led to an excitatory increase followed by an inhibitory decrease in Poisson rate in both the test and flank neurons. The resultant spike trains were cross-correlated and normalized exactly like the real data. For details, see Supplementary Information.

Model of arborization overlap. The basin of common input interactions $F(x, y)$ feeding from points (x, y) on cortex into a neuron at a given recording site (x_0, y_0) was equated to a normalized gaussian weighting function, $F(x, y) = A \exp(-(r/577)^2)$ with $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$ in μm and $A = 1/(\pi(577)^2)$. The spatial scale length of 577 μm was selected so that the amplitude of the gaussian fell to 0.15 times the peak at $r = 800 \mu\text{m}$. The relevant optical image was remapped in terms of orientation relative to the value at (x_0, y_0) and the orthogonal cortical columns were identified. The flank stimulus was assumed to drive cortical columns in an orientation-weighted manner: regions orthogonal $\pm 15^\circ$ at full strength, regions between 15° and 22.5° away from orthogonal at half strength and those between 22.5° and 30° away from orthogonal at one-quarter strength. The corresponding areas of the optical map were integrated and weighted by the orientation-dependent weighting as well as by the gaussian arborization function, to give the final modelled value of arborization overlap with the flank cortical column.

Received 11 January; accepted 26 April 1999.

- Gilbert, C. D. Adult cortical dynamics *Physiol. Rev.* **78**, 467–485 (1998).
- Kapadia, M. K., Ito, M., Gilbert, C. D. & Westheimer, G. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron* **15**, 843–856 (1995).
- Gilbert, C. D. Horizontal integration and cortical dynamics. *Neuron* **9**, 1–13 (1992).
- Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T. & Norcia, A. M. Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature* **391**, 580–584 (1998).
- Westheimer, M. *Laws of Organization in Perceptual Forms* (Harcourt, Brace and Jovanovich, London, 1938).
- Ullman, S. Three-dimensional object recognition. *Cold Spring Harbor Symp. Quant. Biol.* **55**, 889–898 (1990).
- Field, D. J., Hayes, A. & Hess, R. F. Contour integration by the human visual systems: evidence for a local ‘association field’. *Vision Res.* **33**, 173–193 (1993).
- Hubel, D. H. & Wiesel, T. N. Sequence regularity and geometry of orientation columns in the monkey striate cortex. *J. Comp. Neurol.* **158**, 267–294 (1974).
- Ts'o, D. Y., Frostig, R. D., Lieke, E. E. & Grinvald, A. Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science* **249**, 417–420 (1990).

- Blasdel, G. G. & Salama, G. Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* **321**, 579–585 (1986).
- Bonhoeffer, T. & Grinvald, A. Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature* **353**, 429–431 (1991).
- Das, A. & Gilbert, C. D. Distortions of visuotopic map match orientation singularities in primary visual cortex. *Nature* **387**, 594–598 (1997).
- Malach, R., Amir, Y., Harel, M. & Grinvald, A. Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proc. Natl Acad. Sci. USA* **90**, 10469–10473 (1993).
- Malach, R. Dendritic sampling across processing streams in monkey striate cortex. *J. Comp. Neurol.* **315**, 303–312 (1992).
- Hubener, M. & Bolz, J. Relationships between dendritic morphology and cytochrome oxidase compartments in monkey striate cortex. *J. Comp. Neurol.* **324**, 67–80 (1992).
- Katz, L. C., Gilbert, C. D. & Wiesel, T. N. Local circuits and ocular dominance columns in monkey striate cortex. *J. Neurosci.* **9**, 1389–1399 (1989).
- Bosking, W. H., Zhang, Y., Schofield, B. & Fitzpatrick, D. Orientation selectivity and arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* **17**, 2112–2127 (1997).
- Sillito, A. M., Grieve, K. L., Jones, H. L., Cudiero, J. & Davis, J. Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* **378**, 492–496 (1995).
- Shevelev, I. A., Novikova, R. V., Lazareva, N. A., Tikhomirov, A. S. & Sharaev, G. A. Sensitivity to cross-like figures in the cat striate neurons. *Neuroscience* **69**, 51–57 (1995).
- Shevelev, I. A., Lazareva, N. A., Sharaev, G. A., Novikova, R. V. & Tikhomirov, A. S. Selective and invariant sensitivity to crosses and corners in cat striate cortex neurons. *Neuroscience* **84**, 713–721 (1998).
- Gilbert, C. D. & Wiesel, T. N. Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. *Nature* **280**, 120–125 (1979).
- Gilbert, C. D. & Wiesel, T. N. Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* **3**, 1116–1133 (1983).
- Ts'o, D. Y., Gilbert, C. D. & Wiesel, T. N. Relationships between horizontal corticocortical connections in cat visual cortex. *J. Neurosci.* **9**, 2432–2442 (1989).
- Gilbert, C. D. & Wiesel, T. N. Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci.* **9**, 2432–2442 (1989).
- Grinvald, A., Lieke, E. E., Frostig, R. D., Gilbert, C. D. & Wiesel, T. N. Functional architecture of cortex revealed by optical imaging of intrinsic signals. *Nature* **324**, 361–364 (1986).
- Frostig, R. D., Lieke, E. E., Ts'o, D. Y. & Grinvald, A. Cortical functional architecture and local coupling between neuronal activity and the microcirculation revealed by in vivo high resolution optical imaging of intrinsic signals. *Proc. Natl Acad. Sci. USA* **6082**–6086 (1990).
- Das, A. & Gilbert, C. D. Receptive field expansion in adult visual cortex is linked to dynamic changes in strength of cortical connections. *J. Neurophysiol.* **74**, 779–792 (1995).
- Aertsen, A. M. H. J. & Gerstein, G. L. Evaluation of neuronal connectivity: sensitivity of cross-correlation. *Brain Res.* **340**, 341–354 (1985).
- Perkel, D. H., Gerstein, G. L. & Moore, G. P. Neuronal spike trains and stochastic point processes. I. The single spike train. *Biophys. J.* **7**, 391–418 (1967).
- Perkel, D. H., Gerstein, G. L. & Moore, G. P. Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys. J.* **7**, 419–440 (1967).
- Melssen, W. J. & Epping, W. J. M. Detection and estimation of neuronal connectivity based on crosscorrelation analysis. *Biol. Cybern.* **57**, 403–414 (1987).
- Hata, Y., Tsumoto, T., Sato, H. & Tamura, H. Horizontal interactions between visual cortical neurones studied by cross-correlation analysis in the cat. *J. Physiol.* **441**, 593–614 (1991).
- Kisvárdy, Z. F. et al. Synaptic targets of HRP-filled layer III pyramidal cells in the cat striate cortex. *Exp. Brain Res.* **64**, 541–552 (1986).
- McGuire, B. A., Gilbert, C. D., Rivlin, P. K. & Wiesel, T. N. Targets of horizontal connections in macaque primary visual cortex. *J. Comp. Neurol.* **305**, 370–392 (1991).
- Hirsch, J. & Gilbert, C. D. Synaptic physiology of horizontal connections in the cat's visual cortex. *J. Neurosci.* **11**, 1800–1908 (1991).
- Malach, R. Cortical columns as devices for maximizing neuronal diversity. *Trends Neurosci.* **17**, 101–104 (1994).
- Bauer, U., Scholz, M., Levitt, J. B., Obermayer, K. & Lund, J. S. A model for the depth-dependence of receptive field size and contrast sensitivity of cells in layer 4C of macaque striate cortex. *Vision Res.* **39**, 613–629 (1999).
- Hubel, D. H. & Wiesel, T. N. Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B* **198**, 1–59 (1977).
- Ito, M. & Gilbert, C. D. Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* **22**, 593–604 (1999).
- Crist, R. E., Ito, M., Westheimer, G. & Gilbert, C. D. Task dependent contextual interactions in the primary visual cortex of primates trained in hyperacuity discrimination. *Abstr. Soc. Neurosci.* **23**, 1543 (1997).
- Ito, M., Westheimer, G. & Gilbert, C. D. Attention and perceptual learning modulate contextual influences on visual perception. *Neuron* **20**, 1191–1197 (1998).
- Sompolinsky, H. & Shapley, R. New perspectives on the mechanism for orientation selectivity. *Curr. Opin. Neurobiol.* **7**, 514–522 (1997).
- Crook, J. M., Kisvárdy, Z. F. & Eysel, U. T. GABA-induced inactivation of functionally characterized sites in cat striate cortex: effects on orientation tuning and direction selectivity. *Vis. Neurosci.* **14**, 141–158 (1997).
- Crook, J. M., Kisvárdy, Z. F. & Eysel, U. T. Evidence for a contribution of lateral inhibition to orientation tuning and direction selectivity in cat visual cortex: reversible inactivation of functionally characterized sites combined with neuroanatomical tracing techniques. *Eur. J. Neurosci.* **10**, 2056–2075 (1998).
- Gawne, T. J., Kjaer, T. W., Hertz, J. A. & Richmond, B. J. Adjacent visual cortical complex cells share about 20% of their stimulus-related information. *Cerebral Cortex* **6**, 482–489 (1996).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank A. Glatz and J. Lopez for expert technical assistance. The work was supported by the National Science Foundation.

Correspondence and requests for materials should be addressed to C.D.G. (e-mail: gilbert@rockvax.rockefeller.edu).

Identification of two sources of carbon monoxide in comet Hale–Bopp

Michael A. DiSanti^{*,†}, Michael J. Mumma[†],
Neil Dello Russo^{*,†}, Karen Magee-Sauer[‡], Robert Novak[§]
& Terrence W. Rettig^{||}

^{*} Department of Physics, The Catholic University of America,
Washington DC 20064, USA

[†] Laboratory for Extraterrestrial Physics, NASA Goddard Space Flight Center,
Code 690, Greenbelt, Maryland 20771, USA

[‡] Department of Chemistry and Physics, Rowan University, Glassboro,
New Jersey 08028, USA

[§] Department of Physics, Iona College, New Rochelle, New York 10801, USA

^{||} Department of Physics and Astronomy, University of Notre Dame, Notre Dame,
Indiana 46556, USA

The composition of ices in comets may reflect that of the molecular cloud in which the Sun formed, or it may show evidence of chemical processing in the pre-planetary accretion disk around the proto-Sun. As carbon monoxide (CO) is ubiquitous in molecular clouds^{1,2}, its abundance with respect to water could help to determine the degree to which pre-cometary material was processed, although variations in CO abundance may also be influenced by the distance from the Sun at which comets formed^{3–5}. Observations have not hitherto provided an unambiguous measure of CO in the cometary ice (native CO). Evidence for an extended source of CO associated with comet Halley was provided by the Giotto spacecraft^{6–9}, but alternative interpretations exist¹⁰. Here we report observations of comet Hale–Bopp which show that about half of the CO in the comet comes directly from ice stored in the nucleus. The abundance of this CO with respect to water (12 per cent) is smaller than in quiescent regions of molecular clouds, but is consistent with that measured in proto-stellar envelopes¹¹, suggesting that the ices underwent some processing before their inclusion into Hale–Bopp. The remaining CO arises in the coma, probably through thermal destruction of more complex molecules.

The appearance of comet Hale–Bopp enabled studies of the evolution of cometary activity in unprecedented detail. At large

heliocentric distances ($R_h \approx 4–7$ AU), volatile release was driven by CO sublimation^{12–14}. However, gas production rates obtained at large R_h do not provide information on the amount of native CO relative to H₂O, the dominant native ice, due to the substantially different volatilities of these two compounds. Closer to the Sun, water sublimation controls the release of all species, and reliable mixing ratios for native ices can be obtained. Our study of water, CO and dust reveals the onset and persistence of a distributed source for CO within 2 AU heliocentric distance, and establishes the relative abundances of water and CO ices in the cometary nucleus.

We observed Hale–Bopp using the cryogenic echelle grating spectrometer (CSHELL)¹⁵ on the NASA IRTF 3-m telescope at Mauna Kea Observatory, Hawaii (Table 1). Cometary line and continuum intensities were obtained from two-dimensional frames (256 spectral pixels (columns) by ~ 150 spatial pixels (rows)) representing the net (sky-subtracted) comet signal (Fig. 1). Absolute flux calibration was achieved through comparison with the spectra of flux-standard stars. Details of our observing strategy and data processing are given elsewhere^{16,17}. The intensity of CO line emission at each point along the slit is proportional to the column density of molecules in the upper state of the transition giving rise to the line. The CO rotational temperature (T_{rot}) was obtained at each position from multiple lines in the $v = 1–0$ band (where v is the vibrational quantum number), using the standard analysis for linear molecules¹⁸ (Table 1).

A one-dimensional distribution of CO molecules about the nucleus was obtained by combining spatial profiles for several CO lines; dust profiles were obtained from simultaneously measured thermal continuum emission (Fig. 2). The spatial distributions of CO and dust were similar when the comet was beyond ~ 2 AU from the Sun, but the CO emission consistently displayed a more extended profile within ~ 1.5 AU of the Sun. (An extended CO distribution has also been reported from independent CSHELL observations obtained on UT 1997 March 05 ($R_h = 1.03$ AU pre-perihelion).¹⁹) If CO and dust were released only from the nucleus, they should have similar spatial profiles over this region of the coma. The much broader extent for CO that first appears at $R_h \approx 1.5$ AU suggests the onset of a second (distributed) source.

Evidence of both native and distributed CO contributions can be seen by examining the effective production rate (Q , molecules s^{-1}) as a function of projected distance from the nucleus (Fig. 3). Production rate curves ('Q-curves') are generated by stepping a 1×1 arcsec aperture along the CO profile, extracting a production

Table 1 CO production in comet Hale–Bopp

UT date	R_h^* (AU)	Δ^* (AU)	$T_{\text{rot}}^\dagger$ (K)	$Q_{\text{CO}}^\ddagger$ (10^{29} molecules s^{-1})	native/total [§]	$Q_{\text{H}_2\text{O}}^\parallel$ (10^{30} molecules s^{-1})	$Q_{\text{CO}}/Q_{\text{H}_2\text{O}}^\parallel$	$Q_{\text{nat}}/Q_{\text{H}_2\text{O}}^\#$
1996								
13 Jun.	4.11	3.18	20	0.96 ± 0.15				
19 Sep.	3.03	2.92	30	1.28 ± 0.26				
11 Dec.	2.02	2.83	55 ± 9	2.98 ± 0.32				
1997								
21 Jan.	1.48	2.18	94 ± 5	10.78 ± 0.30	0.49	4.00 ± 0.39	0.270 ± 0.027	0.132 ± 0.013
24 Feb.	1.11	1.57	114 ± 6	18.16 ± 0.75	0.48	8.89 ± 0.89	0.204 ± 0.022	0.098 ± 0.011
02 Mar.	1.06	1.47	111 ± 3	19.80 ± 0.57	0.48	7.63 ± 0.23	0.260 ± 0.011	0.125 ± 0.005
10 Apr.	0.93	1.44	105 ± 4	22.31 ± 1.35	0.53	9.99 ± 0.38	0.223 ± 0.016	0.118 ± 0.008
16 Apr.	0.95	1.53	116 ± 4	24.69 ± 2.04	0.46	9.12 ± 0.36	0.271 ± 0.025	0.125 ± 0.011
30 Apr.	1.05	1.75	106 ± 3	14.27 ± 1.00	0.67	6.76 ± 0.33	0.211 ± 0.018	0.141 ± 0.012
01 May.	1.06	1.77	97 ± 3	17.93 ± 0.87	0.53	7.43 ± 0.41	0.241 ± 0.018	0.128 ± 0.009
08 Aug.	2.24	2.90	54 ± 12	2.95 ± 0.37				
25 Sep.	3.08	2.86	30	1.52 ± 0.35				

* Heliocentric and geocentric distances.

† Derived CO rotational temperatures, with uncertainties representing 1σ confidence levels. The rotational temperature increases when moving away from the nucleus, and this is probably the signature of photolytic heating in the intermediate coma, caused primarily by (fast) hydrogen atoms, and/or possibly by electron collisions²⁵. Rotational temperatures were determined from line intensities over a range of projected distances from the nucleus corresponding to the 'terminal' portion of the CO Q-curve. For the three dates with $R_h > 3$ AU, T_{rot} was adopted from ref. 26, and an uncertainty of ± 5 K has been included in calculating Q_{CO} for these dates.

‡ Total (native + distributed) CO production rate. For each date, Q was calculated line-by-line using g -factors appropriate to the derived rotational temperature. Q_{CO} represents the average for all lines observed, and listed uncertainties represent 1σ stochastic errors and include line-to-line variations. These incorporate 14–20 points per line, and multiple lines per date (compare Fig. 3). As we rely on terminal Q-curve values, this method is independent of possible effects in the inner coma (for example, CO optical depth). Further refinements require modelling the distributed CO source (see text).

§ Amount of total CO as ice native to the nucleus (based on values from the 'native CO' heliocentric fit extrapolated to smaller R_h ; compare Fig. 4), divided by Q_{CO} .

|| Derived water production rate¹⁷, with 1σ errors determined as above (footnote †).

Total Q_{CO} relative to $Q_{\text{H}_2\text{O}}$. The mean value for these dates was 0.241 ± 0.009 .

Inferred 'native CO' relative to H₂O. The mean value for these dates was 0.124 ± 0.004 . Thus the mean fractional abundance of 'native CO' to total CO was $51 \pm 3\%$.

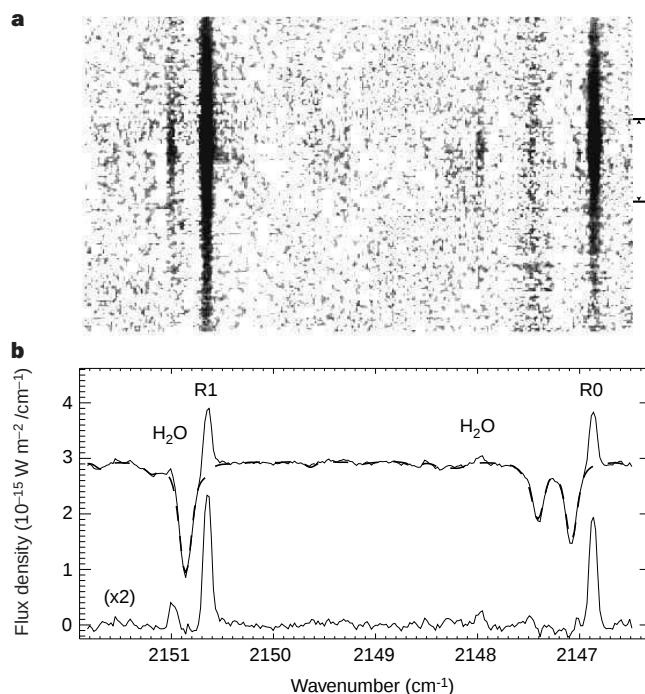


Figure 1 CSHELL spectrum of comet Hale-Bopp, obtained on 1997 May 01.1. **a**, Sky-subtracted spectral-spatial image. To reveal better the distribution of CO emission along slit, the (extremely bright) dust continuum emission has been removed. The small pixel size (0.2 arcsec) afforded by CSHELL provided seeing-limited (frequently 1 arcsec or better) spatial resolution along the 30-arcsec-long slit, which we oriented east-west. The greater spatial extent of the CO lines, compared with that of the adjacent H₂O emissions (marked in **b**), is evident. The vertical scale to the right of the figure indicates 10⁴ km at the comet, and east is at the top. **b**, Top trace: spectral extract (before continuum subtraction) summed over 5 rows (1 arcsec) centred on the peak continuum intensity. Subtraction of the optimized atmospheric model (dashed curve), generated using the Spectrum

Synthesis program²⁷ which accesses the HITRAN-1992 molecular data base²⁸, yields the observed net cometary molecular emission in excess of the continuum (bottom trace; scaled by 2x). We used a 1-arcsec slit width, resulting in a spectral resolving power $\nu/\Delta\nu \approx 2 \times 10^4$, where ν is the photon wavenumber (cm⁻¹). This was sufficient to isolate line emission in the CO $\nu = 1-0$ band (near 4.7 μ m) from adjacent CO absorption features in the terrestrial atmosphere, and from underlying cometary continuum emission. The telescope was periodically moved 2 arcmin north or south of the comet to sample background sky emission; no cometary CO emission was present in the 'sky' beam at the noise level of the data, based on test observations.

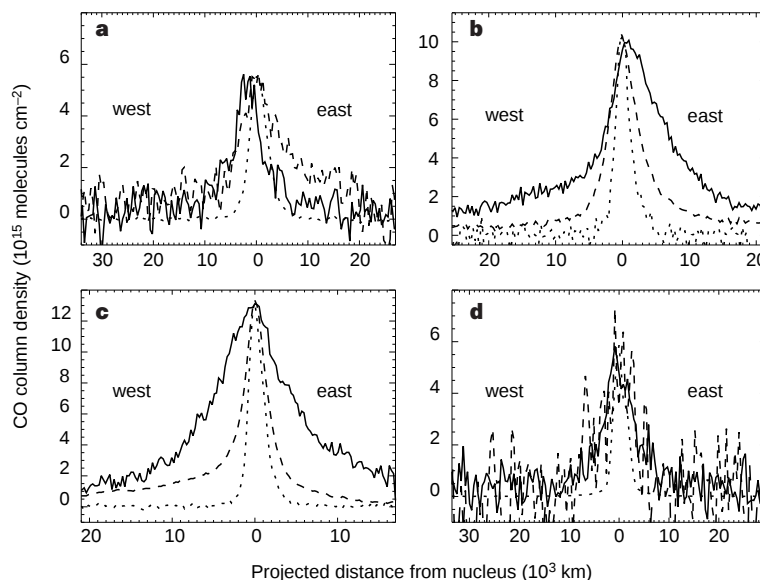


Figure 2 Spatial profiles of CO emission in comet Hale-Bopp. Shown in each panel is the sum of line profiles sampling a range in rotational levels, thereby providing a measure of the overall CO distribution along the slit. **a**, 1996 December 11 ($R_h = 2.02$ AU pre-perihelion); **b**, 1997 January 21 ($R_h = 1.48$ AU pre-perihelion); **c**, 1997 May 01 ($R_h = 1.06$ AU post-perihelion); **d**, 1997 August 08 ($R_h = 2.24$ AU post-perihelion). Scaled continuum (dashed trace) and stellar ('point spread function', dotted trace) profiles are shown for comparison. Note the change in appearance between December and January CO profiles, which we take to signal the onset of the distributed source between 2.0 and 1.5 AU pre-perihelion. At $R_h = 2.02$ AU pre-

perihelion, the spatial profile of CO is east-west symmetric, but the dust is more extended towards the east (**a**). At 1.48 AU and 1.06 AU, the CO profile is markedly more extended than the dust in both east and west directions (**b**, **c**). At 2.24 AU post-perihelion, both CO and dust are east-west symmetric and their profiles have the same extent and shape (**d**). The east-west slit orientation was chosen to facilitate peaking up the comet signal in the slit, and for more easily assessing fine adjustments to telescope tracking rates as needed. The spatial profiles of gas and dust are typically averaged over 2-3 h.

rate at each position, and then averaging values east and west of the nucleus¹⁶. We take the terminal value of the Q -curve to represent the total production rate Q_{CO} (see Table 1). Analogous curves are generated for dust by stepping along the corresponding continuum profile.

Our model assumes constant outflow velocity, so acceleration in the intermediate coma²⁰ could influence the shape of our Q -curves. However, the Q -curves for species released at the nucleus should be similarly affected. Those for water and dust are nearly identical in shape, but differ greatly from that of CO inside 2 AU heliocentric distance (Fig. 3b). Although acceleration introduces small increases to our retrieved gas production rates, another cause must be sought for the strong difference between the Q -curves of water and CO. Giotto measured $\sim 25\%$ increase in gas outflow velocity between $\sim 4 \times 10^3$ km and 2×10^4 km from the nucleus of comet Halley²⁰. A similar effect in Hale-Bopp would lead to enhanced estimates in Q_{CO} (and $Q_{\text{H}_2\text{O}}$) of only 3–4% at 1.48 AU, and 6–8% near perihelion,

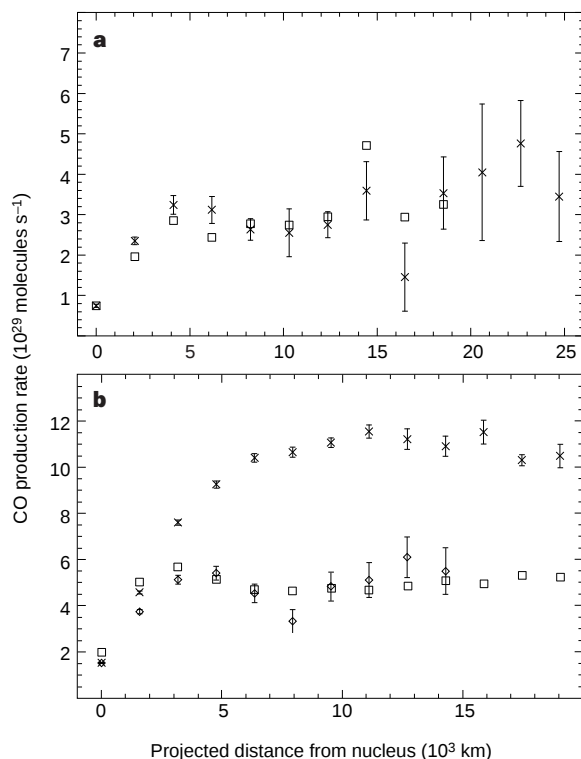


Figure 3 Hale-Bopp production rate curves (Q -curves), showing the derived Q (molecules s^{-1}) as a function of projected distance from the nucleus. The production rate $Q = (F/g)(4\pi\Delta^2/hc\nu\tau)[1/f(x)]$, where F is the emitted line flux (W m^{-2}), g is the fluorescence efficiency (photon s^{-1} molecule $^{-1}$) at $R_h = 1$ AU (Q depends on T_{rot} through g ; ref. 29), Δ is the geocentric distance (m), $hc\nu$ is the energy (J) of a photon with wavenumber ν (cm^{-1}), and τ (s) is the photodissociation lifetime at 1 AU. Under quiet Sun conditions, $\tau_{\text{CO}} \approx 1.2 \times 10^6$ s (ref. 30). The quantity $f(x)$ represents the fraction of molecules included in the sampled region, given square pixels (compare ref. 16). We adopt a model in which production occurs only at the nucleus, and assume a uniform outflow velocity (v) of $1,100R_h^{0.5} \text{ m s}^{-1}$ (compare refs 13, 26). ($Q \propto v$ as $f(x) \propto 1/v$; see text for the estimated effect of gas acceleration on Q .) The average of Q east and west of the nucleus is shown (crosses, with 1σ uncertainties, denote CO), using a 1×1 arcsec stepped aperture. The corresponding continuum Q s (open squares) are shown for comparison. **a**, 1996 December 11 ($R_h = 2.02$ AU pre-perihelion); **b**, 1997 January 21 ($R_h = 1.48$ AU pre-perihelion). For December, Q_{CO} is based on four lines, and Q was averaged 2–11 arcsec either side of the nucleus; a purely native source is indicated (compare Fig. 2a). For January, ten lines were measured and Q was averaged 4–11 arcsec from the nucleus. Terminal values of scaled continuum (or H_2O , open diamonds) and CO Q -curves are in the ratio 0.49 (compare Table 1 footnote §, and Fig. 4). Figure 3 suggests a significant contribution from the distributed CO component in January which was not present in December.

far smaller than the $\sim 100\%$ increase in Q_{CO} over $Q_{\text{H}_2\text{O}}$ seen in Fig. 3b. The presence of a distributed source for CO, similar to that inferred for comet Halley⁶, is suggested.

The evolution from a purely native source for CO to a combination of native and distributed sources occurs between 2.02 and 1.48 AU pre-perihelion; this is clearly seen by comparing Q -curves for these dates. When release is solely from the nucleus, the Q -curve reaches its terminal value quickly—in this case the curves for dust and CO have similar shapes (Fig. 3a). But when both native and distributed sources are present, as for CO, the Q -curve reaches its terminal value farther from the nucleus (Fig. 3b). The ratio of contributions from native and distributed sources may be estimated by comparing the terminal value of the Q -curve for CO with that for dust and H_2O , these latter Q -curves being proxies for the native source alone.

The dependence of total CO production rate on heliocentric distance (Fig. 4) can also be used to distinguish native from distributed sources. Our observations at $R_h > 2$ AU are consistent with $Q_{\text{CO,nat}} = (1.05 \times 10^{30})R_h^{-1.73 \pm 0.26}$ molecules s^{-1} , which we take to represent the (purely) native contribution to total CO based on the similar spatial extent for CO and continuum profiles. This heliocentric dependence is consistent with that derived for H_2O using the same instrument and approach; our Q -curves for water support its release primarily as a parent volatile from the nucleus¹⁷.

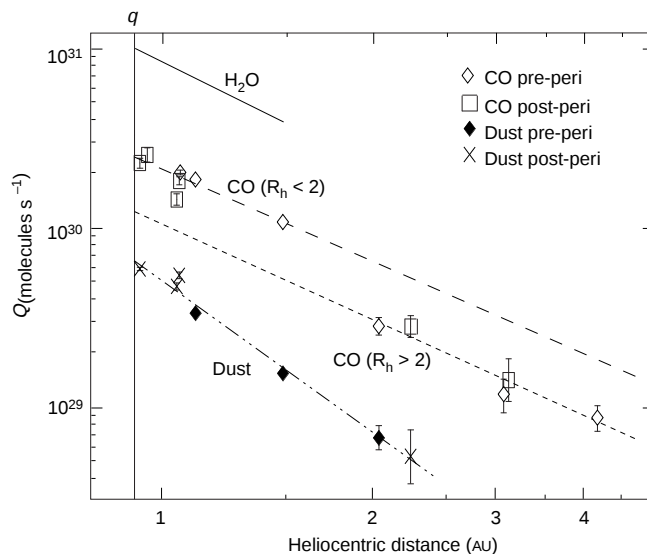


Figure 4 Heliocentric dependence of Q_{CO} and $Q_{\text{H}_2\text{O}}$ in comet Hale-Bopp. The pre-perihelion increase in CO production between 2.0 and 1.5 AU is evident. A least squares linear fit for $R_h > 2$ AU represents the contribution from native CO (short-dashed line) and has a power-law slope of -1.73 ± 0.26 . A fit for $R_h < 2$ AU (long-dashed line) reveals a similar slope (-1.66 ± 0.22), but is displaced upwards by a factor of ~ 2 . These fits imply values of $(1.05 \pm 0.44) \times 10^{30}$ and $(2.07 \pm 0.20) \times 10^{30}$ molecules s^{-1} , respectively, at $R_h = 1$ AU. A fit to $Q_{\text{H}_2\text{O}}$ (solid line) obeys a power law with slope -1.88 ± 0.18 (ref. 17). Relative dust production (Q_{dust}) was obtained from continuum flux density F_c ($\text{W m}^{-2} \text{cm}^{-1}$) after distinguishing contributions from grain thermal emission [$F_c(\text{th})$] and scattered sunlight, using observations obtained near 2.0, 3.35 and 4.7 μm . Q_{dust} was calculated as $F_c(\text{th})/B_\nu(T_{\text{dust}})$, averaged over CO settings. The Planck function $B_\nu(T) \propto \nu^3/[\exp(h\nu/kT) - 1]$, where $x = hc/k \approx 1.44 \text{ cm K}$, and $T = T_{\text{dust}} = 329R_h^{-0.53}$ (ref. 31). $B_\nu(T_{\text{dust}})$ was normalized to unity for $\nu = 2,150 \text{ cm}^{-1}$ (4.651 μm) and $R_h = 1$ AU. Q_{dust} obeys a steeper power law than CO, H_2O , or mm-sized dust³² (-2.73 ± 0.12 , assuming that dust and gas outflow velocities vary similarly with R_h) and exhibits no such discontinuity between 2.0 and 1.5 AU as seen for Q_{CO} . This figure suggests that sublimation of native CO and release of the 'parent' species for distributed CO are controlled by sublimation of the same nuclear ice, and that the onset of distributed CO is not explained in terms of increased dust production or dust fragmentation alone. The vertical solid line (indicated by q) marks the perihelion distance (0.914 AU).

A separate fit for $R_h < 2$ AU ($Q_{CO} = (2.07 \times 10^{30}) R_h^{-1.66 \pm 0.22}$) shows the same heliocentric dependence, and is displaced above the 'native' fit by a factor of ~ 2 . Thus the present analysis suggests approximately equal contributions from native and distributed sources to the total CO production in comet Hale-Bopp within $R_h \sim 1.5$ AU (Table 1).

Could the onset of a distributed source for CO be related to a change in dust production or destruction? Our measurements of dust (at $4.7 \mu\text{m}$ wavelength) sample particles with sizes of a few micrometres, and follow a heliocentric dependence ($Q_{\text{dust}} \propto R_h^{-2.73 \pm 0.12}$) which is steeper than our findings for Q_{CO} and Q_{H_2O} (see note added in proof, below). This implies a rapid increase in dust fragmentation, and hence a decrease in mean particle size, with decreasing heliocentric distance. (In our analysis, we have assumed the same heliocentric dependence of outflow velocity ($R_h^{-0.5}$) for dust and gas.) Detailed modelling of dust properties (outflow, composition, destruction) convolved with geometrical effects, will be reported elsewhere.

If the onset of distributed CO emission was related to increased dust release as the comet approached perihelion, a sharp increase (that is, a 'jump') in Q_{dust} similar to that observed for Q_{CO} between 2.02 and 1.48 AU would be expected. Whereas Q_{CO} follows two distinct distributions—one for $R_h < 2$ AU and another for $R_h > 2$ AU— Q_{dust} appears to fit a single distribution, with no such jump near (or inside) $R_h = 2$ AU.

The lack of a jump in Q_{dust} near $R_h = 2.0$ – 1.5 AU demonstrates that the onset of distributed CO emission cannot be explained by increased dust production alone. Rather, it is plausible that a threshold (most likely thermal) is reached for release of additional CO in the coma. The similar power-law dependence exhibited by Q_{CO} for both $R_h < 2$ AU and $R_h > 2$ AU, and by Q_{H_2O} (ref. 17), suggests that release of the 'parent' giving rise to distributed CO is controlled by sublimation of the same nuclear ice as seen for other native volatiles (water, native CO). We thus infer that the distributed parent has a constant mixing ratio in the nucleus relative to native H_2O and CO ices.

The difference between the Q-curves for CO and (scaled) dust approximates the Q-curve for the distributed component alone. This reaches its terminal value at a projected distance $\rho \approx 6.5 \times 10^3$ km from the nucleus for $R_h = 1.48$ AU, and at $\sim 5.1 \times 10^3$ km for $R_h = 1.06$ AU. The distance ρ approximates the scale for release of distributed CO, and can be compared with model predictions to better constrain the scale length of the distributed parent. Future refinements^{21–24} to our analytical approach include accounting for axisymmetric outflow and for gas acceleration in the coma. We expect such modelling to have relatively little effect on our retrieved amount of native CO, but it should provide significant improvements to our understanding of the distributed source.

We have characterized the release of carbon monoxide in comet Hale-Bopp, and compared it with water and dust. Within $R_h \approx 1.5$ AU, our results support a dual-source nature for CO production, as seen first in comet Halley during the Giotto spacecraft encounter⁶. We infer a total (native plus distributed) mean CO production rate of $24.1 \pm 0.9\%$ relative to water, with $51 \pm 3\%$ of the observed CO contained in the nuclear ice, the remainder being produced in the coma.

Note added in proof: Production of millimetre-sized dust in comet Hale-Bopp follows a single distribution with power law $R_h^{-1.7 \pm 0.2}$ for R_h between 0.9 and 2.5 AU (ref. 32). This heliocentric slope is consistent with those we derive for Q_{CO} and Q_{H_2O} , and this supports uniform mixing of volatile and refractory components in the nucleus. □

Received 17 November 1998; accepted 12 April 1999.

- Rank, D. M., Townes, C. H. & Welch, W. J. Interstellar molecules and dense clouds. *Science* **174**, 1083–1101 (1971).
- Turner, B. E. Recent progress in astrochemistry. *Space Sci. Rev.* **51**, 235–337 (1989).
- Mumma, M. J. Organic volatiles in comets: Their relation to interstellar ices and solar nebula material. *Astron. Soc. Pacif. Conf. Ser.* **122**, 369–396 (1997).
- Mumma, M. J., Weissman, P. R. & Stern, S. A. in *Protostars and Planets, III* (eds Levy, E. H. & Lunine, J. I.) 1177–1252 (Univ. Arizona Press, Tucson, 1993).

- Sandford, S. A. & Allamandola, L. J. The condensation and vaporization behavior of $H_2O:CO$ ices and implications for interstellar grains and cometary activity. *Icarus* **76**, 201–224 (1988).
- Eberhardt, P. *et al.* The CO and N_2 abundance in comet P/Halley. *Astron. Astrophys.* **187**, 481–484 (1987).
- Huebner, W. F., Boice, D. C. & Sharp, C. M. Polyoxymethylene in comet Halley. *Astrophys. J.* **320**, L149–L152 (1987).
- Huntress, W. T., Allen, M. & Delitsky, M. Carbon suboxide in comet Halley? *Nature* **352**, 316–318 (1991).
- Meier, R., Eberhardt, P., Krankowsky, D. & Hodges, R. R. The extended formaldehyde source in comet P/Halley. *Astron. Astrophys.* **277**, 677–690 (1993).
- Greenberg, J. M. & Li, A. From interstellar dust to comets: the extended CO source in comet Halley. *Astron. Astrophys.* **332**, 374–384 (1998).
- Chiar, J. E. *et al.* Processing of icy mantles in protostellar envelopes. *Astrophys. J.* **498**, 716–727 (1998).
- Biver, N. *et al.* Substantial outgassing of CO from comet Hale-Bopp at large heliocentric distance. *Nature* **380**, 137–139 (1996).
- Biver, N. *et al.* Evolution of the outgassing of comet Hale-Bopp (C/1995 O1) from radio observations. *Science* **275**, 1915–1918 (1997).
- Womack, M., Festou, M. C. & Stern, S. A. The heliocentric evolution of key species in the distantly-active comet C/1995 O1 (Hale-Bopp). *Astron. J.* **114**, 2789–2795 (1997).
- Greene, T. P., Tokunaga, A. T., Toomey, D. W. & Carr, J. S. CSHELL: A high spectral resolution 1–5 μm cryogenic echelle spectrograph for the IRTF. *Proc. SPIE* **1946**, 311–324 (1993).
- Dello Russo, N., DiSanti, M. A., Mumma, M. J., Magee-Sauer, K. & Rettig, T. W. Carbonyl sulfide in comets C/1996 B2 (Hyakutake) and C/1995 O1 (Hale-Bopp): Evidence for an extended source in Hale-Bopp. *Icarus* **135**, 377–388 (1998).
- Dello Russo, N. *et al.* Direct detection of water in comet C/1995 O1 (Hale-Bopp). *Icarus* (submitted).
- Herzberg, G. *Spectra of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1950).
- Weaver, H. A. *et al.* Infrared spectroscopy of comet Hale-Bopp. *Earth Moon Planets* (in the press).
- Lämmerzahl, P. *et al.* Expansion velocity and temperatures of gas and ions measured in the coma of comet P/Halley. *Astron. Astrophys.* **187**, 169–173 (1987).
- Boice, D. C., Sablik, M. J. & Konno, I. Distributed coma sources and the CH_4/CO ratio in Comet Halley. *Geophys. Res. Lett.* **17**, 1813–1816 (1990).
- Crifo, J. F. A general physicochemical model of the inner coma of active comets. 1. Implications of spatially distributed gas and dust production. *Astrophys. J.* **445**, 470–488 (1995).
- Xie, X. & Mumma, M. J. Monte Carlo simulation of cometary atmospheres: Application to Comet P/Halley at the time of the Giotto Spacecraft encounter. II. Axisymmetric model. *Astrophys. J.* **464**, 457–475 (1996).
- Combi, M. R. Time-dependent gas kinetics in tenuous planetary atmospheres: The cometary coma. *Icarus* **123**, 207–226 (1996).
- Xie, X. & Mumma, M. J. The effect of electron collisions on rotational populations of cometary water. *Astrophys. J.* **386**, 720–728 (1992).
- Biver, N. *et al.* Long term evolution of the outgassing of comet Hale-Bopp from radio observations. *Earth Moon Planets* (in the press).
- Kunde, V. G. & Maguire, W. C. A direct integration transmittance model. *J. Quant. Spectrosc. Rad. Transf.* **14**, 803–817 (1974).
- Rothman, L. S. *et al.* The HITRAN molecular database: Editions of 1991 and 1992. *J. Quant. Spectrosc. Rad. Transf.* **48**, 469–507 (1992).
- Crovisier, J. Rotational and vibrational synthetic spectra of linear parent molecules in comets. *Astron. Astrophys. Suppl.* **68**, 223–258 (1987).
- Huebner, W. F., Keady, J. J. & Lyon, S. P. Solar photo rates for planetary atmospheres and atmospheric pollutants. *Astrophys. Space Sci.* **195**, 1–294 (1992).
- Green, S. F., McDonnell, J. A. M., Pankiewicz, G. S. A. & Zarnecki, J. C. in *20th ESLAB Symposium on the Exploration of Halley's Comet Vol. 2*, 81–86 (SP-250, ESA, 1986).
- Jewitt, D. & Matthews, H. E. Particulate mass loss from comet Hale-Bopp. *Astron. J.* **117**, 1056–1062 (1999).

Acknowledgements. This work was supported through the NASA Planetary Astronomy Program. We thank the staff of the NASA Infrared Telescope Facility for their support throughout our comet Hale-Bopp observing campaign. The IRTF is operated by the University of Hawaii under contract to NASA. We thank J. Crovisier for comments which improved the manuscript.

Correspondence and requests for materials should be addressed to M.D. (e-mail: disanti@kuiper.gsfc.nasa.gov).

Collective and plastic vortex motion in superconductors at high flux densities

A. M. Troyanovskii[†], J. Aarts^{*} & P. H. Kes^{*}

^{*}Kamerlingh Onnes Laboratory, Leiden University, PO Box 9504, 2300 RA Leiden, The Netherlands

[†]Institute for High Pressure Physics, Russian Academy of Science, Troitsk, 142092, Russia

The 'mixed state' of type II superconductors occurs when magnetic flux penetrates the material (in the form of vortices) without destroying the superconducting ground state. Zero resistivity is retained if the vortices are pinned by crystalline defects, but is destroyed by vortex motion. This provides the practical motivation for studying vortices in random pinning potentials^{1–7}. But the insights so obtained also bear on the more general class of problems involving the dynamics of elastic media in the presence

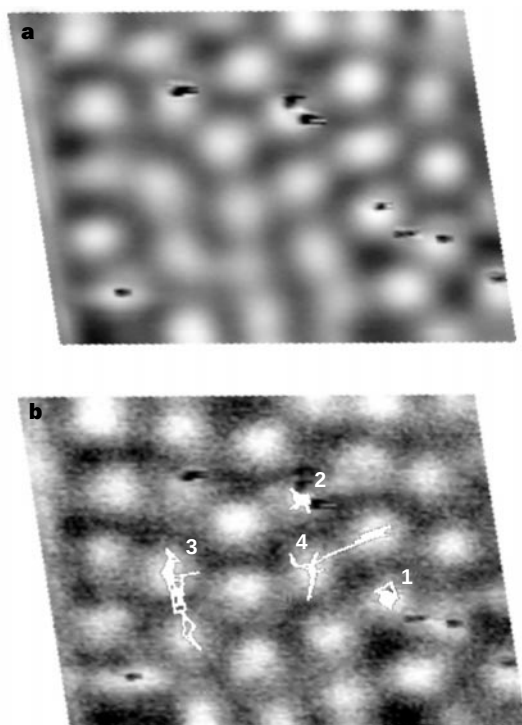


Figure 1 STM image (420×280 nm) at $B = 0.6$ T of plastic flow of the vortex lattice in NbSe₂ containing columnar defects (CDs). The CDs are seen as black dots, their concentration is $B_0 = 0.3$ T (at B_0 the areal density of vortices equals that of the CDs). Panel **a** shows an average over 128 subsequent images, resulting in a grey scale representing the probability density for finding a vortex at a certain position. Several trajectories of plastic flow can be recognized. Panel **b** shows the trajectories of four typical vortices as obtained from images 0–35. See text for details of these vortices.

of quenched disorder⁸ (for example, mechanical friction). Moreover, the magnetic vortex system is highly tunable and permits questions concerning frictional, plastic and elastic flow⁹ to be investigated on the scale of single vortices. Remarkable results have been obtained on the dynamics of this system^{10–17}, but have been largely restricted to well separated vortices at very low flux densities. Scanning tunnelling microscopy has the potential to resolve individual vortices at much higher flux densities^{18–21}, and here we show that the imaging rates can be sufficiently high to resolve the dynamics in this flux regime. We find that, in the presence of strongly pinning line defects, the vortex lattice remains pinned until the number of vortices is about twice that of the defects, at which point plastic creep commences. But in the presence of weak intrinsic point disorder, the vortices creep coherently along one of the principle axes of the vortex lattice, where they exhibit striking and unanticipated velocity modulations that appear to be related to the lattice periodicity.

Our scanning tunnelling microscope (STM) experiments were all performed at 4.3 K in fields B up to 1.9 T. From SQUID (superconducting quantum interference device) magnetization measurements on a pristine NbSe₂ crystal, we determined a critical temperature $T_c = 7.10$ K and an upper critical field $\mu_0 H_{c2}(4.3 \text{ K}) = 2.13$ T. The coherence length ξ at 4.3 K is therefore 12.5 nm and the vortex core diameter about 35 nm (ref. 22). Several crystals were irradiated with 6-GeV Pb ions in order to create columnar defects (CDs). Two doses were used, corresponding to B_0 of 0.2 and 0.3 T (at B_0 the areal density of vortices equals that of the CDs). The typical sample size was about 2×2 mm and the thickness ranged between 0.1 and 0.5 mm. The thinnest crystals were used for the ion irradiation to ensure that the CDs traverse the entire

sample. Before the experiment, the sample was mounted onto the sample holder and cleaved under ambient conditions. Our STM is designed as an insert to a helium transport vessel, similar to that described in ref. 23. The STM with sample was cooled down to 4.3 K in zero applied field by slowly immersing it into liquid helium. A superconducting coil was used to generate a magnetic field in the direction of the CDs and perpendicular to the sample surface.

The STM tip was positioned somewhere between the edge and the centre of the sample. Vortices were imaged using a cut Pt–Ir tip both in topographic mode (constant current) and current mode (constant average current). They appear, depending on B , as hills of height 0.1–2 Å because of the modulation of the tunnelling conductance associated with the order-parameter variations of the vortex lattice. With increasing B , the amplitude of the modulation decreases and the images become more noisy. Tunnelling currents between 15 and 25 pA and a tip–sample bias voltage of 0.5–1 mV were used. For a scanning area of 400×400 nm with 128 lines the scanning time was ≥ 3 s. Data storage took 3.3 s so that the total acquisition time of one image could be as small as 6.5 s. Usually it was ~ 10 s. This fast scanning operation allowed us to observe moving vortices by STM compiled in a cartoon-like fashion as vortex movies; see Supplementary Information.

The raw data seem to show a somewhat distorted vortex lattice; but a hexagonal vortex arrangement results after a correction is made for the spatial distortions of our microscope in the x – y plane. The scaling factors were determined by comparing the image of the NbSe₂ lattice with the crystallographic atom arrangement. The field calibration was carried out at fields close to H_{c2} where the demagnetization correction is negligible. The position of the CDs in irradiated samples provided information about the mechanical stability of the STM and sample holder. As there was no visible displacement of CDs in the images, we estimate the instability to be less than 1–2 pixels.

The STM images contain both topographic and spectroscopic information, and it is possible to see CDs and vortices simultaneously. The defects are visible as black spots (see Fig. 1) with diameters that appear to be 10–20 nm, but this is an artefact of our experiment; it is known³⁰ that the track diameter is about 7.0–9.0 nm. The distribution of defects is random with large fluctuations, which explains why in Fig. 1 we only observe half the average number (~ 16) expected for this area. Because of the competition between pinning energy and deformation energy, the occupation number at $B \approx B_0$ is only 50% (ref. 24). For the same reason the vortex arrangement at $B < B_0$ is very disordered²⁵. At fields far above B_0 the vortex configurations approaches that of a granular triangular lattice with (strong) perturbations at the sites of the CDs.

To study vortex motion in the sample with $B_0 = 0.3$ T, we collected data consecutively at fields of 0.6, 0.5, 0.4, 0.3 and 0.2 T after coming down from 1.0 T and waiting 20 min each time before making measurements. The same area, shown in Fig. 1, was repeatedly probed. The vortex lattice at all fields was highly disordered. For fields up to 0.5 T, essentially all vortices remain pinned during the observation time (500 s). Even a disordered vortex lattice does not lose its integrity when only $\sim 1/3$ of its vortices are trapped on a CD, revealing that the effective shear modulus is still appreciable. At 0.6 T plastic flow sets in, as is seen in Fig. 1a which shows an average over 128 subsequent images. The grey scale thus represents the probability of finding a vortex at a certain position. In the lower left corner is one of the flow paths of vortices moving through more strongly pinned neighbours. In Fig. 1b we show the vortex motion by marking the subsequent positions (for frames 0–35) of four typical vortices: vortex 1, pinned by 1 CD; vortex 2, pinned by 2 CDs; vortex 3 in a path of plastic flow; and vortex 4, which occasionally slips through between vortices 1 and 2. We note that ‘vortex 3’ in fact represents three different vortices which subsequently move along the same path. When we concentrate on vortex 1, we see that it moves isotropically in the

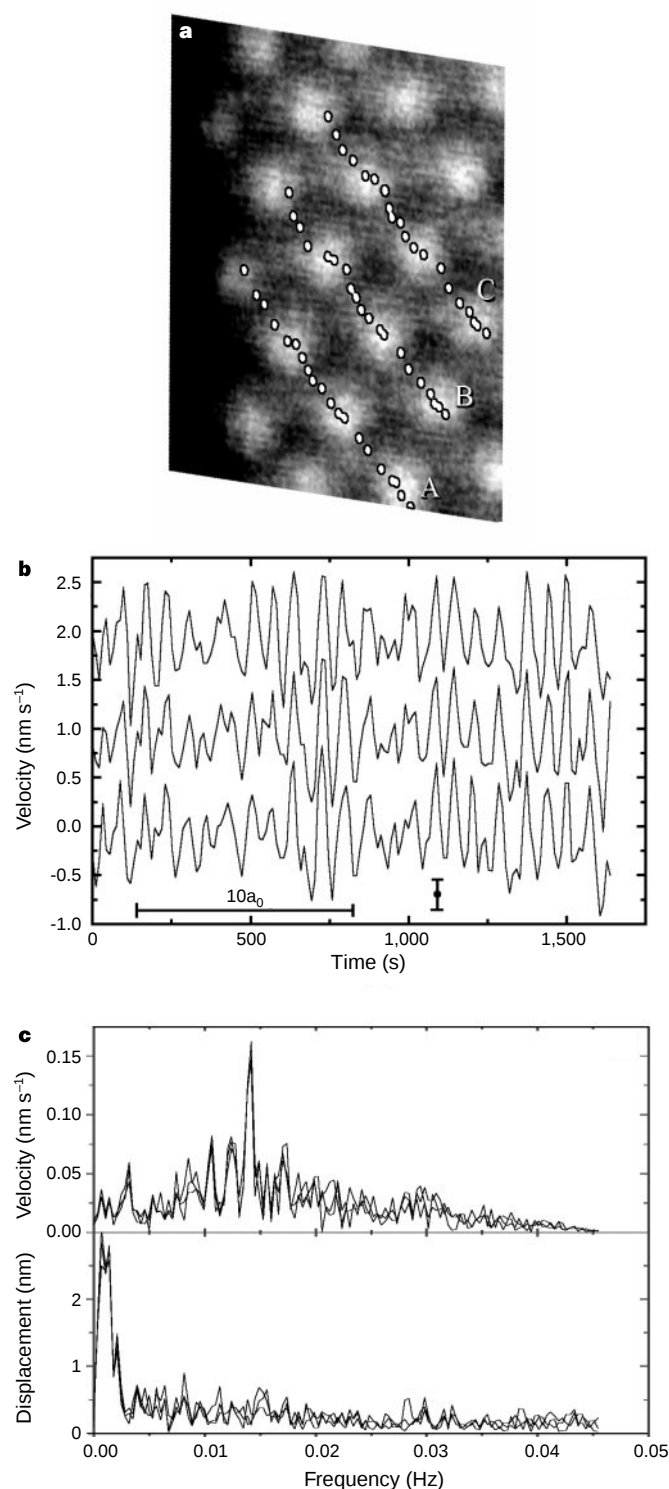


Figure 2 STM image (230×355 nm) of a moving vortex lattice in pristine NbSe₂ at a field 0.6 T. **a**, Frame no. 85 out of a sequence of 168 images. The dots represent the subsequent positions of three vortices (A, B, C) as determined from frames nos 84–103. Imaging time for each frame is 6.5 s, repetition time is 11 s. **b**, The longitudinal velocity of the three vortices in **a** versus time (data for vortices A and C are shifted over ± 1.0 nm s⁻¹ for clarity). A velocity modulation with period a_0 (the vortex lattice parameter) can be clearly identified. **c**, Fourier transforms of the longitudinal velocity (after subtracting the average value of v to suppress the peak at zero frequency) and the transverse displacements of the same vortices A, B and C. The peak at 14.1 mHz in the upper panel reveals the “washboard” frequency (see text) at v/a_0 for the longitudinal velocity and displacements. No peak is seen for the transverse displacements in the lower panel. Here the low-frequency peaks are displaying the long-timescale (700–1,500 s) fluctuations.

potential well caused by the CD as it is pushed around by its moving neighbours. The motion of vortex 2 is more anisotropic, with a larger amplitude parallel to the line connecting the two CDs on which it is pinned. The average distance between vortex 1 and 2 is 115 nm; this is slightly smaller than $2a_0$ ($= 126$ nm), where $a_0 = (2\Phi_0/\sqrt{3}B)^{1/2}$ is the vortex lattice parameter and Φ_0 the flux quantum. This prohibits the movement of vortex 4. However, the displacement fluctuations caused by the other moving vortices may add roughly 13 nm to the average distance, allowing vortex 4 to slip through and follow a path that seems to merge with that of vortex 3.

For the pristine NbSe₂ crystals, almost perfect hexagonal vortex lattices were observed in a wide range of magnetic fields. Figure 2a shows a typical image: it was taken 20 min after the field was increased from 0 to 0.6 T and the persistent mode switch turned on. It contains 22 vortices, and is one of 168 subsequent images each taking 11 s yielding a total recording time of 30 min. It is remarkable that the 22 vortices move coherently as one bundle. Twenty subsequent positions of three vortices (A, B, C) are denoted by the white dots. Within experimental error they each represent the motion of the entire bundle. We convinced ourselves by decreasing the scan rate that we are not observing some stroboscopic effect but that we are really detecting the average velocity v . Its mean value is about $a_0/6$ per frame, yielding $v \approx 1.0$ nm s⁻¹ and an electric field $E = vB \approx 0.6$ nV m⁻¹. This value is 5 orders of magnitude smaller than the commonly used voltage criterion defining the critical current density j_c . The motion can be self-consistently²⁶ described as collective creep with a resistivity $\rho \approx \rho_f \exp[-U(j)/k_B T]$ where ρ_f is the flux flow resistivity. The actual energy barrier $U(j)$ and current density j turn out to be $j \approx 0.4j_c \approx 2.3 \times 10^6$ A m⁻² and $U \approx 77$ K. The transverse size R_A of the region in which positional short-range order persists in absence of topological defects can be estimated, assuming collective pinning²⁷, to be $R_A \approx R_c(a_0/\xi\sqrt{2})^{4-d}$, where R_c is the transverse Larkin length²⁷ and d the dimensionality of the vortex lattice disorder. With $R_c \approx 3a_0$ (ref. 28) we obtain $R_A \approx 36a_0$ for $d = 3$, or $R_A \approx 10a_0$ for $d = 2$. Our observation of a perfect vortex lattice consisting of 22 vortices falls well within this estimate, irrespective of d . Finally, it is seen that the mean distance the vortex bundle has moved between frames is ~ 12 nm, which is quite close to the value of ξ at 4.3 K. This observation fits with the picture of creep in which vortex bundles hop between local minima in the random pinning potential which has the coherence length as the shortest relevant length scale.

The vortex lattice is seen (Fig. 2a) to move along one of its principle axes. This seems to be in agreement with the minimum dissipation argument of Schmid and Hauger²⁹ for a vortex lattice moving with large velocity (flow regime) through a random potential. However, it is unlikely that this argument should hold for the creep regime as well. On the other hand, in equilibrium, the orientations of the vortex lattice and the underlying crystal lattice of NbSe₂ coincide¹⁸ because both have hexagonal symmetry. This coincidence, which we indeed observe, will impose an additional ‘tin roof’ potential on the moving vortex lattice thus fixing its orientation, but still allowing sliding motion when the driving force f_d is not aligned with the principle axis. We do not know the direction of f_d , but it is quite unlikely that it would have been directed along a principle axis of the vortex lattice in all our experiments. Nevertheless, we do not see sliding motion. Our data therefore rather support the predictions of Le Doussal and Giamarchi^{3,4} that a slowly moving glass at finite temperatures behaves like a moving glass at $T = 0$, that is, it moves along rough channels (see below) which are stable against a small transverse force component.

Our most important result follows from plotting the positions of the vortices A, B and C versus time. When they moved out of the field of view we added na_0 , with integer n , to the position of the corresponding vortex that just appeared in the picture. This procedure is allowed because the vortices move coherently with

well known average velocity ($v = 0.89 \text{ nm s}^{-1}$) and orientation. Taking the derivative of these trajectories yields a plot of the longitudinal velocities versus time (Fig. 2b). For clarity, we shifted the results for vortex A and C by $\pm 1 \text{ nm s}^{-1}$. The error bar denotes the uncertainty in v ($\delta v = \pm 0.15 \text{ nm s}^{-1}$) due to the error in the vortex positions. The horizontal bar marks the time needed to traverse a distance $10a_0$ (630 nm). It is clearly seen that the velocity is modulated with a periodic component with amplitude $\Delta v \approx 0.6 \text{ nm s}^{-1}$ and period a_0/v . In the Fourier transform (Fig. 2c), this shows up as a peak at 14.1 mHz. Such a peak at the “washboard” frequency is known for periodic structures driven through random point disorder⁹, but to the best of our knowledge this is the first time that this washboard feature has been found to survive in the creep regime. The periodic modulation also appears in the longitudinal displacement determined with respect to the average, uniform motion versus time. However, it is not observed in a plot of the transverse displacements versus time. Rather, we see on our timescale a random-walk-like transverse excursion of the bundle up to an amplitude $0.2a_0$, not yet enough to decide how the mean-squared displacement grows with time. We leave this for future investigations.

Fast imaging of vortex lattices by STM provides the possibility of studying the collective and plastic flow behaviour of elastic media through various configurations of disorder. □

Received 7 December 1998; accepted 5 April 1999.

- Blatter, G. *et al.* Vortices in high-temperature superconductors. *Rev. Mod. Phys.* **66**, 1125–1388 (1994).
- Koshlev, A. E. & Vinokur, V. M. Dynamic melting of the vortex lattice. *Phys. Rev. Lett.* **73**, 3580–3583 (1994).
- Giamarchi, T. & Le Doussal, P. Moving glass phase of driven lattices. *Phys. Rev. Lett.* **76**, 3408–3411 (1996).
- Le Doussal, P. & Giamarchi, T. Moving glass theory of driven lattices with disorder. *Phys. Rev. B* **57**, 11356–11403 (1998).
- Moon, K., Scalettar, R. & Zimányi, G. T. Dynamical phases of driven vortex systems. *Phys. Rev. Lett.* **77**, 2778–2781 (1996).
- Scheidt, S. & Vinokur, V. M. Driven dynamics of periodic elastic media in disorder. *Phys. Rev. B* **57**, 13800–13810 (1998).
- Balents, L., Marchetti, M. C. & Radziko, L. Nonequilibrium steady states of driven periodic media. *Phys. Rev. B* **57**, 7705–7739 (1998).
- Nattermann, T. Scaling approach to pinning: Charge-density waves and giant flux creep in superconductors. *Phys. Rev. Lett.* **64**, 2454–2457 (1990).
- Olsen, C. J., Reichhardt, C. & Nori, F. Nonequilibrium dynamics phase diagram for vortex lattices. *Phys. Rev. Lett.* **81**, 3757–3760 (1998).
- Osakabe, N., Kasai, H., Kodama, T. & Tonomura, A. Time-resolved analysis in transmission electron microscopy and its application to the study of the dynamics of vortices. *Phys. Rev. Lett.* **78**, 1711–1714 (1997).
- Kirtley, J. R. *et al.* Direct imaging of integer and half-integer Josephson vortices in high- T_c grain boundaries. *Phys. Rev. Lett.* **76**, 1336–1339 (1996).
- Oral, A. *et al.* Direct observation of melting of the vortex solid in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ single crystals. *Phys. Rev. Lett.* **80**, 3610–3613 (1998).
- Mozer, A. *et al.* Observation of single vortices condensed into a vortex-glass phase by magnetic force microscopy. *Phys. Rev. Lett.* **74**, 1847–1850 (1995).
- Bolle, C. A., de la Cruz, F., Gammel, P. L., Waszczak, J. V. & Bishop, D. J. Observation of tilt induced orientational order in the magnetic flux lattice 2H-NbSe_2 . *Phys. Rev. Lett.* **71**, 4039–4042 (1993).
- Yao, Z. *et al.* Path of magnetic flux lines through high- T_c copper oxide superconductors. *Nature* **371**, 777–779 (1994).
- Marchevsky, M. V., Aarts, J., Kes, P. H. & Indenbom, M. V. Observation of the correlated vortex flow in NbSe_2 . *Phys. Rev. Lett.* **78**, 531–534 (1997).
- Pardo, F. *et al.* Real space images of the vortex lattice structure in a Type II superconductor during creep over a barrier. *Phys. Rev. Lett.* **79**, 1369–1372 (1997).
- Hess, H. F., Robinson, R. B. & Waszczak, J. V. Vortex-core structure observed with a scanning tunneling microscope. *Phys. Rev. Lett.* **64**, 2711–2714 (1990).
- Maggio-Aprile, I., Renner, Ch., Erb, A., Walker, E. & Fischer, Ø. Direct vortex lattice imaging and tunneling spectroscopy of flux lines on $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. *Phys. Rev. Lett.* **75**, 2754–2757 (1995).
- Renner, Ch., Revaz, B., Kadowaki, K., Maggio-Aprile, I. & Fischer, Ø. Observation of the low temperature pseudogap in the vortex cores of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Phys. Rev. Lett.* **80**, 3606–3609 (1998).
- De Wilde, Y. *et al.* Scanning tunneling microscopy observation of a square Abrikosov lattice in $\text{LuNi}_2\text{B}_2\text{C}$. *Phys. Rev. Lett.* **78**, 4273–4276 (1997).
- Volodin, A. P., Golubov, A. A. & Aarts, J. Vortex core shapes measured by STM. *Z. Phys. B* **102**, 317–321 (1997).
- Volodin, A. P. & Troyanovskii, A. M. A three-phase piezoinertio motor for a low temperature STM. *Instr. Exp. Techn.* **40**, 724–726 (1997).
- Wengel, C. & Tauber, U. C. Weakly pinned Bose glass vs Mott insulator phase in superconductors. *Phys. Rev. Lett.* **78**, 4845–4848 (1997).
- Behler, S. *et al.* Vortex pinning in ion-irradiated NbSe_2 studied by scanning tunneling microscopy. *Phys. Rev. Lett.* **72**, 1750–1753 (1994).
- van der Beek, C. J., Nieuwenhuys, G. J. & Kes, P. H. Nonlinear current diffusion in type-II superconductors. *Physica C* **197**, 320–336 (1992).
- Larkin, A. I. & Ovchinnikov, Yu. N. Pinning in type II superconductors. *J. Low Temp. Phys.* **34**, 409–427 (1979).
- Agurel, L. A., Amin, F., Polichetti, M., Aarts, J. & Kes, P. H. Dimensionality of collective pinning in 2H-NbSe_2 single crystals. *Phys. Rev. B* **56**, 3425–3432 (1997).
- Schmid, A. & Hauger, W. On the theory of vortex motion in an inhomogeneous superconducting film. *J. Low Temp. Phys.* **11**, 667–685 (1973).

30. Bauer, P. *et al.* Depth sensitive visualisation of irradiation-induced columnar defects in 2H-NbSe_2 . *Euro. Phys. Lett.* **23**, 585–591 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>).

Acknowledgements. We thank A. P. Volodin for developing the first version of our STM, R. J. Drost for auxiliary experiments, M. Konczykowski for carrying out the heavy ion irradiation at GANIL (Caen, France), and M. V. Marchevsky, P. Le Doussal and T. Giamarchi for discussions. This project was partially supported by FOM and the Dutch-Russian Science Collaboration financed by NWO.

Correspondence and requests for materials should be addressed to P.H.K. (e-mail: kes@phys.leidenuniv.nl).

Full sintering of powdered-metal bodies in a microwave field

Rustum Roy, Dinesh Agrawal, Jiping Cheng & Shalva Gedevanishvili

Materials Research Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

The use of microwaves to process absorbing materials was studied intensively in the 1970s and 1980s, and has now been applied to a wide variety of materials^{1–4}. Initially, success in microwave heating and sintering was confined mainly to oxide and some non-oxide ceramics^{5–11}; but recently the technique has been extended to carbide semimetals^{12–14} used in cutting tools. Here we describe the microwave sintering of powdered metals to full density. We are able to sinter a wide range of standard powdered metals from commercial sources using a 2.45-GHz microwave field, yielding dense products with better mechanical properties than those obtained by conventional heating. These findings are surprising in view of the reflectivity of bulk metals at microwave frequencies. The ability to sinter metals with microwaves should assist in the preparation of high-performance metal parts needed in many industries, for example, in the automotive industry.

Reviews of microwave processing^{1–5} describe its use for materials ranging from wood, bacon and potato chips to rubber, ceramics and semiconductors, but make no mention of metal sintering. Walkiewicz *et al.*¹⁵ exposed a range of materials, including six metals (presumably partly oxidized in air) to a 2.4-GHz field, and reported modest heating (but not sintering), ranging from 120 °C for Mg to 768 °C for Fe. Sintering of tungsten carbide–cobalt composites has also been reported¹⁶. Nishitani reported¹⁷ that by adding a few per cent of electrically conducting powders such as aluminium, the heating rates of refractory ceramics was considerably enhanced; but no mention was made of the microwave sintering of pure metal powders. Whittaker and Mingos¹⁸ used the highly exothermic (and rapid) reactions of metal powders with sulphur for the microwave-induced synthesis of metal sulphides. German¹⁹ discusses microwave heating of oxide and non-oxide ceramics, but does not mention sintering of metals.

We have used commercial powdered metal components of various alloy compositions—including iron and steel, copper, aluminium, nickel, molybdenum, cobalt, tungsten, tungsten carbide and tin, and their alloys²⁰—to obtain essentially fully dense bodies with substantially improved mechanical properties compared to similar bodies sintered by conventional thermal means. Some of these samples were obtained from Keystone Powdered Metal Company (St. Mary, Pennsylvania, USA) and some were made in our laboratory. They were all standard ‘green’ parts—metal powders cold-pressed with a few per cent of organic binder. The typical size range was 2–5 cm in the largest dimension, with a wide variety of rectangular and cylindrical cross-sections, including typical toothed gears in the 1–4-cm range.

Over the past decade we have developed several specialized

microwave sintering chambers capable of processing a variety of samples with different shapes and sizes, and maintaining any temperature in the range from room temperature to 2,000 °C. The microwave generators are operated at a frequency of 2.45 GHz, with power output in the range 1–6 kW, in both single- and multi-mode operation. Inside a typical microwave cavity is an alumina tube surrounded by ceramic fibre (typically mullite) insulation. The primary function of insulation is to preserve the heat inside the tube. Insulation does not absorb microwaves in the lower temperature ranges, but at higher temperatures there is some partition of the power dissipation between the insulation and the sample. For higher temperatures (above ~1,400 °C) we use ZrO₂ and even Y₂O₃ insulation, but these are much more expensive, and the partition is of course different. For some runs we utilize a secondary coupler or microwave susceptor rods made of SiC or MoSi₂. Temperatures are read by optical pyrometers and/or sheathed thermocouples placed very close to the surface of the sample. Although relative readings are accurate to ±5 °C, the absolute temperature has not been established to anywhere near that precision. The atmosphere can be controlled as needed: from air, to H₂, to forming gas, to oxygen. A diagram of a typical microwave system for sintering of powdered metals is shown in Fig. 1.

The 'green' commercial powdered-metal bodies are introduced into the microwave chamber and heated at 1,100–1,300 °C typically for times ranging from 5 minutes to one hour in flowing forming gas (N₂ + H₂) or in a pure hydrogen atmosphere. In a large number of cases, similar samples have been heated in conventional furnaces for direct comparison of the properties. The first few runs produced highly sintered bodies in a very short period of time. This was achieved in our controlled-atmosphere microwave system with 2.45-GHz frequency and a maximum 6-kW power: but only 1.4 kW of power was used to attain the desired sintering temperatures. Typically the total cycle time was ~90 minutes. For this sintering process, we have established that similar heating occurs without using susceptors. The only difference is that with susceptors the overall heating is faster. This was proved by conducting several

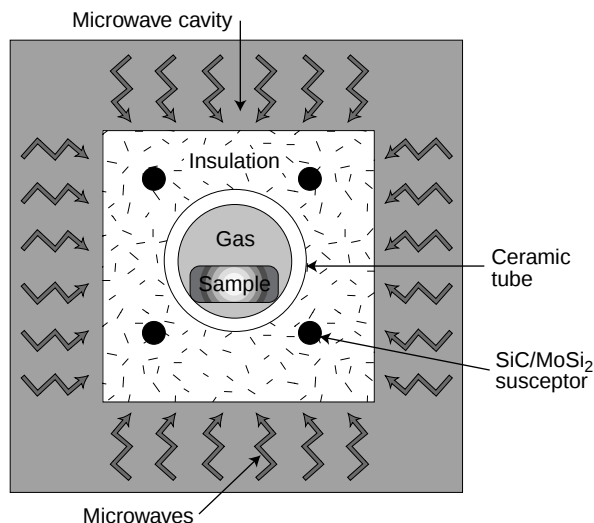


Figure 1 Diagram of microwave system for sintering of powdered-metal parts. The system consists of a 2.54-GHz microwave oven (cavity) and a ceramic (alumina) tube which is inserted into the cavity and surrounded by ceramic fibre blocks. Inside the insulation, SiC or MoSi₂ susceptor rods are inserted. The sample is placed inside the ceramic tube. The system is capable of achieving temperatures up to 1,600 °C, and any desired atmospheres (such as H₂, N₂, Ar) can be used. The same system has been used with or without the susceptor (SiC/MoSi₂) rods; sintering times without the rods are somewhat longer but no other change is noted.

Table 1 Properties of microwave and conventionally processed powdered-metal samples

Sample	Process	Green density (g cm ⁻³)	Sintered density (g cm ⁻³)	Rockwell hardness	MOR (10 ³ lb in ⁻²)
Fe-Ni (industrial part)	MW	7.11	7.15	B82	177
	Conv.		7.10	B77	109
Fe-Cu (industrial part)	MW	6.81	7.17	B96	142
	Conv.		6.84	B80	118
Fe-Cu (lab. sample)	MW	6.95	6.96	B75	134
	Conv.		6.95	B64	122

MW, microwave processed; conv., conventionally processed. The modulus of rupture (MOR) of most microwave-processed samples is higher than that of the conventional samples. The densities of many microwave-processed samples are also higher than those of conventional samples.

experiments by removing the susceptors altogether and monitoring the power level. It was also established in essentially all cases that the net shape of the green part was retained precisely (with the usual shrinkage if any was planned), and a fine microstructure was produced.

Table 1 shows data for some of these microwave-processed parts, and the corresponding properties of parts of the same composition made by conventional means. Our findings indicate that every powdered-metal 'green' body so far tested could be sintered in 10–30 minutes in an appropriate microwave sintering apparatus. It is also clear that in almost all cases, the modulus of rupture (MOR) of microwave-processed samples is substantially higher than the conventional samples; in the case of the Fe–Ni composition, it was 60% higher. The densities of many microwave-processed samples are also higher than conventional samples.

The samples with a composition of Fe + Cu(2%) + graphite (0.8%) were processed in a microwave field at 1,200 °C for 30 minutes. The microwave-sintered and 'green' samples were characterized for their microstructure by optical microscopy and for phase composition by X-ray diffractometry. The microstructure

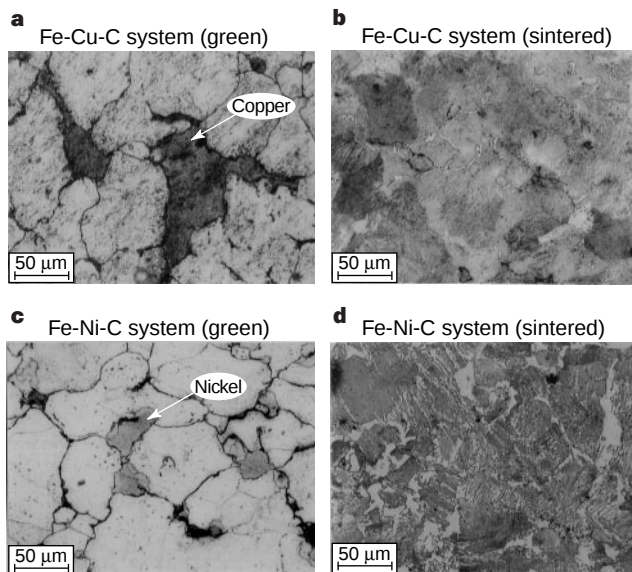


Figure 2 Optical micrographs of two sets of 'green' (that is, unsintered) and microwave-sintered powdered-metal parts processed at 1,200 °C for 30 minutes. **a, b**, Fe-Cu-C; 'green' sample (**a**), microwave-sintered sample (**b**). **c, d**, Fe-Ni-C; 'green' sample (**c**), microwave-sintered sample (**d**). Microwave-sintered powdered-metal parts show excellent sintering of Fe particles. In the Fe-Cu-C system, copper melted and spread into Fe particle boundaries forming Fe-Cu solid solutions (**b**). In the Fe-Ni-C system, the microstructure shows light shades for austenitic nickel-rich islands (**d**). In both cases, the formation of steel was confirmed by X-ray diffraction.

photographs (Fig. 2) indicate that excellent sintering had occurred between the iron particles, and that the copper had melted and spread into the iron-particle boundaries forming Fe–Cu solid solutions. The X-ray diffractograms confirm the microscopy observations, and show that the sintered sample contained only a α -iron solid solution single phase.

We also microwave-sintered several powders of pure metals. Only one representative example is described here. Cobalt metal powder was pressed into pellets, and microwave sintered in pure H_2 at 1 atmosphere pressure at various temperatures ranging from 900 to 1,200 °C for 10 minutes. The density reached 8.70 g cm^{-3} at 900 °C to 8.88 g cm^{-3} at 1,000–1,050 °C, and very near theoretical density of 8.89 g cm^{-3} at 1,100–1,200 °C.

These findings have implications for the field of powdered-metal technology. Metal powders are used in industry for a wide range of products and applications. Traditional powder metallurgy involves a process whereby a metal or alloy powder is compacted as a 'green' body and sintered to net shape at elevated temperatures. This process is not labour-intensive, but does consume considerable energy. It conserves materials, and produces high-quality components with reproducible properties. However, the need for materials of high integrity for advanced engineering applications requires newer technologies. The attainment of finer microstructures and near-theoretical densities in special powdered-metal components is still elusive and challenging. Increasing cost is also a concern. The microwave processing method reported here may offer fine microstructures and better properties in powdered-metal products at lower cost. The main reason why the microwave process yields better mechanical properties is two-fold, especially in the case of powdered metals: it produces finer grain size, and the shape of the porosity (if any) is different from that generated during conventional heating. In microwave-processed powdered-metal samples we observed round-edged porosities, producing higher ductility and toughness.

The implications for the fundamental science of microwave–materials interaction are less clear. What is certain from this work and that of Willert-Porada¹⁴ on ceramic sintering is that these interactions are more complex than hitherto suspected. The following factors, at least, contribute significantly to the total microwave heating of powdered metals (object shape and size are also well-known to play a part, but have not been quantified in detail). First, the microwave field generated by a magnetron and the existence of an automatic cut-off mechanism to avoid burn-out due to reflected power. Second, the presence of SiC susceptor rods—taking up only a small fraction of the total space. Third, the existence of a cylindrical insulation package made of aluminosilicate fibres essentially transparent to 2.45-GHz radiation at room temperature, surrounding an alumina tube. Last, the size and nature of the sample placed in the centre of the alumina tube (see Fig. 1).

How the microwave energy absorption is divided between the susceptors, insulation, and sample is a generic issue for all microwave processing, and we consider it in more detail in Supplementary Information. This analysis shows that there is no complete theory to explain even the simplest case of sintering of a ceramic. Furthermore, Cherradi *et al.*²¹ recently claimed that in most ceramics the dielectric loss mechanism was a minor contribution to the power absorbed compared to the induction losses caused by eddy currents. These authors also attribute the heating—they did not sintering—of metals also to eddy-current losses from the electric field. Their evidence, obtained using samples of different size and shape in different orientations, supports the role of such eddy-current losses as a major contributor to the heating of metals. However, all their experiments were done in a single-mode cavity where the orientation of the electric and magnetic fields can be fixed. All our work was done in multimode cavities where the situation is much less determined; but the role of the electric field cannot be ignored.

Another possible significant contribution to the absorption of microwave power, on which much theoretical work has been done²², is multiple scattering in powdered assemblages; this is especially applicable to ceramics and higher microwave frequencies. □

Received 9 October 1998; accepted 20 April 1999.

- Clark, D. & Sutton, W. H. Microwave processing of materials. *Annu. Rev. Mater. Sci.* **26**, 299–331 (1996).
- Schiffman, R. F. Commercializing microwave systems: Paths to success or failure. *Ceram. Trans.* **59**, 7–17 (1995).
- Katz, J. D. Microwave sintering of ceramics. *Annu. Rev. Mater. Sci.* **22**, 153–170 (1992).
- Sutton, W. Microwave processing of ceramics: an overview. *Mater. Res. Soc. Symp. Proc.* **269**, 3–19 (1992).
- Sutton, W. Microwave processing of ceramic materials. *Am. Ceram. Soc. Bull.* **68**, 376–386 (1989).
- Fang, Y., Agrawal, D. K., Roy, D. M. & Roy, R. Fabrication of transparent hydroxyapatite ceramics by microwave processing. *Mater. Lett.* **23**, 147–151 (1995).
- Agrawal, D. K., Fang, Y., Roy, D. M. & Roy, R. *Mater. Res. Soc. Symp. Proc.* **269**, 231–236 (1992).
- Fang, Y., Agrawal, D. K., Roy, D. M. & Roy, R. Fabrication of porous hydroxyapatite ceramics by microwave processing. *J. Mater. Res.* **7**, 490–494 (1992).
- Fang, Y., Roy, R., Agrawal, D. K. & Roy, D. M. Transparent mullite ceramics from diphasic aerogels by microwave and conventional processing. *Mater. Lett.* **28**, 11–15 (1996).
- Fang, Y., Agrawal, D. K., Roy, D. M. & Roy, R. Microwave sintering of calcium strontium zirconium phosphate ceramics. *Ceram. Trans.* **36**, 109 (1993).
- Fang, Y., Agrawal, D. K., Roy, D. & Roy, R. Microwave sintering of hydroxyapatite-based composites. *Ceram. Trans.* **36**, 397 (1993).
- Roy, R., Agrawal, D., Cheng, J. P. & Mathis, M. Microwave processing: triumph of applications-driven science in WC-composites and ferroic titanates. *Ceram. Trans.* **80**, 3–26 (1997).
- Cheng, J. P., Agrawal, D. K., Komarneni, S., Mathis, M. & Roy, R. Microwave processing of WC-Co composites and ferroic titanates. *Mater. Res. Innov.* **1**, 44–52 (1997).
- Gerdes, T. & Willert-Porada, M. *Mater. Res. Soc. Symp. Proc.* **347**, 531–537 (1996).
- Walkiewicz, J. W., Kazonich, G. & McGill, S. L. Microwave heating characteristics of selected minerals and compounds. *Min. Metall. Processing* **5**, 39–42 (1988).
- Willert-Porada, M., Gerdes, T., Rodiger, K. & Kolaska, H. Einsatz von Mikrowellen zum Sintern pulvermetallurgischer Produkte. *Metall* **50**, 744–752 (1996).
- Nishitani, T. Method for sintering refractories and an apparatus therefor. US Patent No. 4147911 (1979).
- Whittaker, A. G. & Mingos, D. M. Microwave-assisted solid-state reactions involving metal powders. *J. Chem. Soc. Dalton Trans.* 2073–2079 (1995).
- German, R. M. *Sintering Theory and Practice* (Wiley, New York, 1996).
- Gedevanishvili, S., Agrawal, D. & Roy, R. Microwave combustion synthesis and sintering of intermetallics and alloys. *Mater. Sci. Lett.* (submitted).
- Cherradi, A., Desgardin, G., Provost, J. & Raveau, B. Electric magnetic field contributions to the microwave sintering of ceramics. In *Electroceraamics IV* Vol. II, (eds. Wasner, R., Hoffmann, S., Bonnenberg, D. & Hoffmann, C.) 1219–1224 (RWTH, Aachen, 1994).
- Shanker, B. & Lakhtakia, A. Extended Maxwell Garnett formalism for composite adhesives for microwave-assisted adhesion of polymer surfaces. *J. Compos. Mater.* **27**, 1203–1213 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank J. Kosco for providing powdered-metal components. We also thank the approximately 20 corporations that supported the early stages of this work, which is now supported by the Electric Power Research Institute, the Keystone Powder Metal Company and the Ben Franklin Technology Center.

Correspondence and requests for materials should be addressed to D.A. (e-mail: dx44@psu.edu).

Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation

Hyun Joo, Zhanglin Lin & Frances H. Arnold

Division of Chemistry and Chemical Engineering 210-41,
California Institute of Technology, Pasadena, California 91125, USA

Enzyme-based chemical transformations typically proceed with high selectivity under mild conditions, and are becoming increasingly important in the pharmaceutical and chemical industries. Cytochrome P450 monooxygenases (P450s) constitute a large family¹ of enzymes of particular interest in this regard. Their biological functions, such as detoxification of xenobiotics and steroidogenesis^{2–5}, are based on the ability to catalyse the insertion of oxygen into a wide variety of compounds⁶. Such a catalytic transformation might find technological applications in areas ranging from gene therapy and environmental remediation to

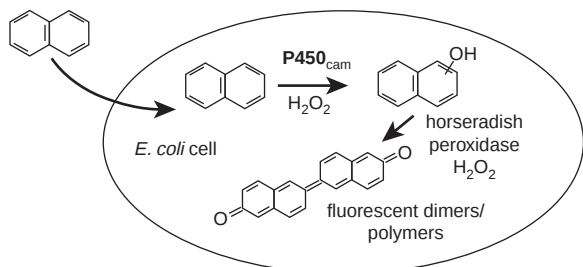


Figure 1 Reaction scheme for detection of active P450_{cam} variants using HRP to generate fluorescent products. The aromatic substrate (here, naphthalene) is taken up by the cells, where it is hydroxylated by the oxygenase. The products of this reaction are oxidatively coupled by HRP, also expressed in the *E. coli*. The product of the coupling reaction is highly fluorescent, and emits at a longer wavelength, relative to the naphthol.

the selective synthesis of pharmaceuticals and chemicals^{7–10}. But relatively low turnover rates (particularly towards non-natural substrates), low stability and the need for electron-donating cofactors prohibit the practical use of P450s as isolated enzymes. Here we report the directed evolution¹¹ of the P450 from *Pseudomonas putida* to create mutants that hydroxylate naphthalene in the absence of cofactors through the ‘peroxide shunt’ pathway^{12,13} with more than 20-fold higher activity than the native enzyme. We are able to screen efficiently for improved mutants by coexpressing them with horseradish peroxidase, which converts the products of the P450 reaction into fluorescent compounds amenable to digital imaging screening. This system should allow us to select and develop mono- and di-oxygenases into practically useful biocatalysts for the hydroxylation of a wide range of aromatic compounds.

Dramatic improvements in cofactor regeneration schemes using a second, coupled enzyme such as formate dehydrogenase have led to the wider use of redox enzymes for chemical synthesis^{14,15}. But coupled enzyme cofactor regeneration is not practical for redox reactions that require multiple proteins (that is, most monooxygenases and dioxygenases). Furthermore, cofactor regeneration requires a second enzyme and a second substrate, and generates a product which must be separated from the desired product. The practical utility of the synthetically versatile P450s would be significantly expanded if we could eliminate the need for a cofactor altogether. Our desire to create practical single-enzyme hydroxylation catalysts led us to attempt to create cytochrome P450s that efficiently utilize hydrogen peroxide (H₂O₂) in lieu of dioxygen and NADH. No longer requiring the P450-associated electron transfer proteins, such catalysts would be extremely simple compared to the natural enzyme systems and could be used in a variety of catalyst forms and reactor configurations.

Horseradish peroxidase (HRP), a highly glycosylated haem peroxidase with four disulphide bridges¹⁶, is not expressed functionally in *Escherichia coli*. However, we generated a variant of HRP that is expressed in active form by directed evolution of the native HRP gene¹⁷. In this variant (HRP1A6), asparagine 255 (a glycosylation site in the native enzyme) is replaced by aspartic acid, a replacement that presumably assists folding in the non-native environment. *E. coli* expressing the monooxygenase cytochrome P450_{cam} from *P. putida*¹⁸ together with HRP1A6 become brightly fluorescent in the presence of naphthalene and H₂O₂ (Figs 1 and 2).

The natural substrate of cytochrome P450_{cam} is camphor; the enzyme shows only weak activity towards naphthalene¹⁹. Approximately 200,000 random mutants of cytochrome P450_{cam} produced by mutagenic polymerase chain reaction (PCR) and coexpressed with HRP1A6 were screened for activity on naphthalene and hydrogen peroxide by fluorescence digital imaging. The fluorescence values of 32,000 clones from this library are plotted in descending order in Fig. 3. Although most of the mutants utilizing

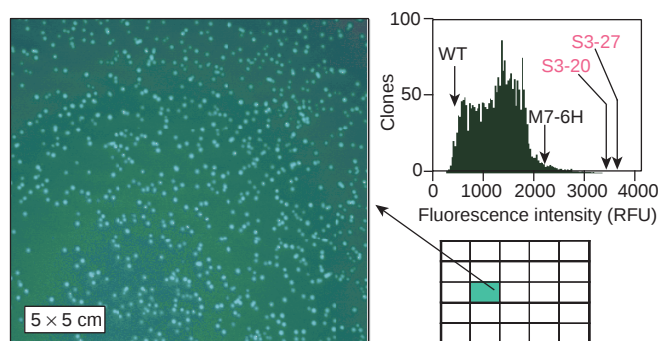


Figure 2 Results of fluorescence measurements. Fluorescence digital (pseudo-colour) image of a portion of a plate with colonies of *E. coli* BL21 (DE3) cells expressing cytochrome P450_{cam} and HRP and containing naphthalene (6 mM) and H₂O₂ (30 mM). Fluorescent products of the HRP coupling reaction are quantified by intensity measurement (see Methods). Histogram on the right side shows the fluorescence values of ~70,000 clones from the library created by StEP recombination of five improved P450_{cam} genes from the first generation (RFU, relative fluorescence units). The fluorescence values of wild-type (WT), improved first-generation variant M7-6H and improved second-generation variants S3-20 and S3-27 are indicated. Large libraries of P450_{cam} variants (here, ~20,000 colonies per plate) can be screened for improved activity, stability or changes in reaction selectivity.

the peroxide shunt pathway are less active than wild type or show no activity at all, a relatively large number show higher levels of fluorescence, indicating higher monooxygenase activity. The fluorescence intensities of control cells either remained constant (host strain with no HRP or P450_{cam}) or showed only a small increase with time (cells expressing the enzymes, but without added naphthalene or H₂O₂, and cells expressing HRP or P450_{cam} singly with naphthalene and H₂O₂; H.J., A. Arisawa, Z.L. and F.H.A., manuscript in preparation).

Three clones with enhanced fluorescence were selected for growth and confirmation of the enhanced activity towards naphthalene in a whole-cell assay. Clone M7-6H showed an 11-fold increase in activity compared to wild-type P450_{cam} (Fig. 3). Two other clones (M7-4H and M7-8H) identified by the digital image scanning also showed at least 5- to 8-fold activity increases. We also assessed the ability of these variants to hydroxylate 3-phenylpropionate (3-PPA) in the whole-cell assay. All three showed improved activities relative to wild-type P450_{cam}; the largest increase was 3.2 fold, for M7-6H. The increased activity of mutant M7-8H was confirmed for the purified enzyme: specific activities of wild-type and M7-8H P450_{cam} were approximately 9 and 70 units (see Methods), respectively, in line with the result from the whole-cell assay.

DNA sequencing revealed the following amino-acid substitutions: E331K (M7-4H); R280L and E331K (M7-6H); E331K and C242F (M7-8H). All three improved P450s have the E331K mutation, which is located in a type I β -turn between the K' helix and a segment often referred to as the ‘meander’^{20,21}. Mutation R280L is in the K helix. Cysteine 242 is at the edge of the proximal haem pocket in the I helix; its replacement with phenylalanine may serve to protect the enzyme from rapid inactivation by hydrogen peroxide²². We can offer no simple rationalization for the R280L and E331K mutations. Near, but not in, the distal (peroxide and ligand binding) pocket, they may serve to improve interactions with either the peroxide or the aromatic substrate. Directed evolution is powerful precisely for this ability to identify mutations that influence enzyme activity through subtle, longer-range interactions¹¹.

In the presence of hydrogen peroxide, HRP catalyses the oxidative coupling of a wide variety of phenols and catechols (>140) to generate coloured and fluorescent materials²⁴. This approach is therefore generally useful for screening both mono- and dioxygenases for their ability to hydroxylate an array of aromatic

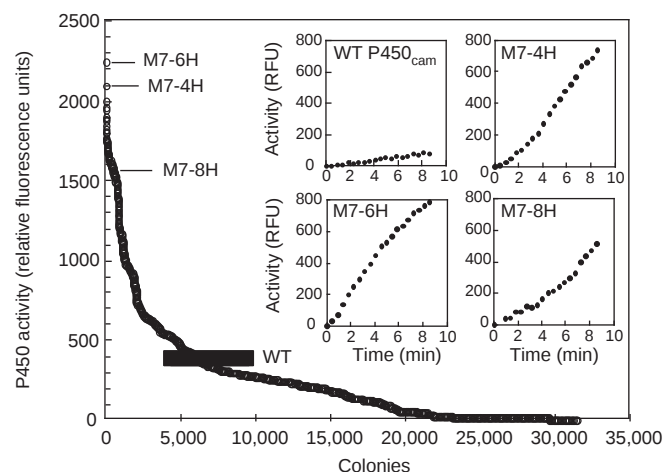


Figure 3 Fluorescence values of *E. coli* clones expressing HRP and random mutants of P450_{cam}, in the presence of naphthalene and H₂O₂. Naphthalene hydroxylation activities, as measured by the total fluorescence generated during a fixed incubation period, are plotted in descending order. Total fluorescence of clones expressing wild-type P450_{cam} is also indicated. Insets, activities of selected clones grown in 50-ml cultures, as measured by fluorescence generation (RFU, relative fluorescence units). Wild-type, 9.2 RFU min⁻¹; M7-4H, 84.1; M7-6H, 86.7; M7-8H, 53.1.

compounds. We expect that some subset of the enzymes that are able to utilize the peroxide shunt pathway efficiently to hydroxylate aromatic substrates will also insert oxygen into other substrates, including alkanes and alkenes. These enzymes could be identified by further screening of the reduced libraries (for example, by liquid chromatography/mass spectrometry, LC/MS). Furthermore, the ultraviolet-visible absorption and fluorescence properties of the products of the HRP reaction can be quite distinct for different hydroxylated products, offering the potential to distinguish oxygenases with different regiospecificities. *In vitro* HRP-catalysed polymerization of various naphthol isomers (1- and 2-) and dihydroxylated naphthalenes (DHN) (1,5-, 1,3-, 2,3- and 2,7-DHN) generates a variety of fluorescent products, with dark blue (430–460 nm), blue-green (495 nm), yellow (580 nm) and orange-red (620 nm) fluorescence. A combination of 1- or 2-naphthol and 2,7-DHN produces a red fluorescent product (620 nm), while combining 1,5-DHN with 2,7-DHN or 2-naphthol produces pink or yellow fluorescence, respectively. 1,3-DHN itself shows pink fluorescence; its polymerization with 1-naphthol increases the intensity and shifts the wavelength to the red. The emission spectra of the products depend on the relative molar ratios of the reactants.

Bacteria expressing wild-type P450_{cam} generate only blue fluorescence (460 nm), corresponding to the conversion of naphthalene to 1- or 2-naphthol. Bacteria expressing the P450_{cam} mutants, however, generate a palette of colours (Fig. 4), reflecting the altered regiospecificities of the P450_{cam}-catalysed hydroxylations. Thin layer chromatography and mass spectrometry performed on the products of the whole-cell reaction showed that wild-type P450_{cam} produces primarily 1-naphthol, with a small amount of 1,3-DHN (1-naphthol:1,3-DHN~4.5:1). In contrast, 'pink' mutant P-67 produces much more of the dihydroxy-compound (1-naphthol:1,3-DHN~1:2). By generating and characterizing P450s with altered substrate and reaction regiospecificities, we should be able to ascertain the molecular determinants of enzyme specificity in this important class of enzymes.

Genes encoding improved P450 variants can be recombined by DNA shuffling methods^{24,25} or further mutated in additional cycles of directed evolution to generate further improved enzymes²⁶. A second-generation P450_{cam} library prepared by staggered extension process (StEP) *in vitro* recombination²⁵ of five improved variants

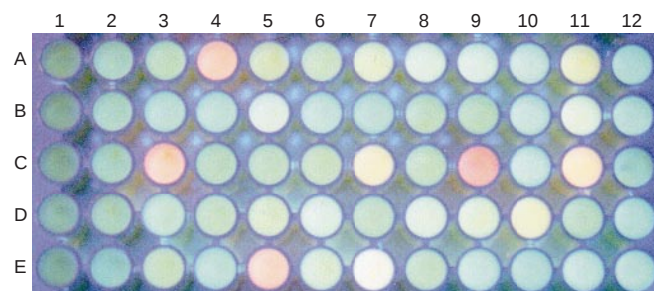


Figure 4 Different colours generated by bacteria expressing HRP and selected P450_{cam} mutants indicating changes in regiospecificity of naphthalene hydroxylation. The first column on the left side (rows A–E) contains control strain *E. coli* BL21 (DE3). The second column contains cells expressing wild-type P450_{cam}. The remaining 50 wells contain different P450_{cam} variants selected by fluorescence image scanning.

has yielded several variants with ~20-fold improvements in naphthalene hydroxylation activity over wild-type. The results of screening this second-generation library are shown in Fig. 2. Two clones isolated from this library are 19- and 21-times more active than wild-type in the whole-cell assay (4 ml culture).

With further evolution of P450s, we would be able to explore the limits of P450-catalysed aromatic hydroxylation via the peroxide shunt pathway. Furthermore, we anticipate that we would be able to identify and evolve thermostable P450s by comparing residual activity after incubation at elevated temperature to initial activity, just as we have done recently to convert a mesophilic esterase²⁷ and protease²⁸ into their thermophilic counterparts. This powerful, simple evolution system should allow us to generate a variety of practical biocatalysts for this useful chemical transformation. It could also be used to screen genetic libraries isolated from the environment for aromatic hydroxylation catalysts.

Cytochrome P450s have presumably evolved by divergent evolution from an ancestral enzyme into a large family of monooxygenases that utilize molecular O₂ as an oxygen source¹. Although the atmosphere of the early Earth contained little or no molecular oxygen, it was relatively rich in H₂O₂ and peroxygenated organic chemicals^{29,30}. The ancient ancestors of today's P450s may have in fact relied on H₂O₂ as the oxygen donor. It is possible that cytochrome P450s that have evolved in the laboratory to function using hydrogen peroxide rather than O₂ could allow us to recreate the functions of these ancestral enzymes. □

Methods

Coexpression of P450_{cam} with HRP. *P. putida* cytochrome P450_{cam} and HRP were coexpressed in *E. coli* BL21(DE3) cells using vectors P450_{cam}-pCWori(+) (provided by P. Ortiz de Montellano) and pETpelBHRP1A6Kan. The plasmid backbone of pCWori(+) contains a double *P*_{tac} promoter and an ampicillin resistance coding region. The gene for HRP1A6 was restricted from pETpelBHRP1A6 (ref. 17) and cloned into the kanamycin resistant vector pET26b(+) (Novagen), yielding pETpelBHRP1A6Kan. Intracellular coexpression of P450_{cam} and HRP1A6 was established on agar/terrific broth (TB) plates supplemented with 0.5 mM isopropyl β-D-thiogalactopyranoside, 0.2 mM thiamine, 1 mM δ-aminolevulinic acid and FeCl₂ trace elements. Cultures were seeded onto agar/TB plates supplemented with 100 μg ml⁻¹ ampicillin and 30 μg ml⁻¹ kanamycin, and were grown at 37°C for 6 h; the incubation temperature was then lowered to 30°C. After 16 h, the colonies were replicated using a nitrocellulose membrane and transferred onto fresh M9 agar (with 10% (w/v) glucose, 5% (v/v) ethanol) plates containing 6 mM naphthalene and 10 mM (first generation) or 30 mM (second generation) hydrogen peroxide for screening by fluorescence image analysis. Cell viability is >99% for a 6 h reaction in the first generation.

Screening by digital imaging. Fluorescence images were digitally scanned (H.J., A. Arisawa, Z.L. and F.H.A., manuscript in preparation) at 460 nm

wavelength using Eagle Eye II and a top-mounted 350-nm ultraviolet illuminator (Stratagene). Automated colony detection and individual fluorescence intensity measurements were performed using Optimas 5.0 image analysis software (Optimas) with a weighted score of 27,000 grey scale. A blue band-pass filter (430–470 nm), 4× lens zoom level, and 1/10 s CCD exposure time were used. Typical colony diameters were ~0.4–0.8 mm. Scanned images were further processed by configuring overall thresholding, geometry recognition, intensity quantification, global and local segmentation, and cutting edges to reduce background fluorescence.

Hydroxylation activity measurements. Naphthalene and 3-PPA whole-cell hydroxylation activities were measured in 200-μl reactions in a 96-well microtitre plate. P450_{cam} + HRP cells grown in 50-ml flasks were collected by centrifugation (Beckman CS SR) at 3,350 r.p.m. and resuspended in 1 ml of 0.1 M sodium phosphate buffer (pH 9.0). A 50-μl aliquot of the cell solution was added to the reaction mixtures (total 200 μl) containing 25% ethanol, 3-PPA (0.5 mM) or naphthalene (6 mM), and H₂O₂ (10 mM) in the same buffer. Fluorescence was measured in a 96-well microfluorimeter (HTS 7000, Perkin Elmer).

For regioselectivity comparisons, reactions were carried out using cells expressing P450_{cam} and no HRP. Cells grown in 50-ml flasks were collected by centrifugation and gently resuspended in 5 ml of 0.1 M phosphate buffer (pH 9.0) containing 5% (v/v) ethanol, 50 mM naphthalene and 50 mM H₂O₂. After 1 h incubation, the products were extracted with 2 ml of chloroform, and the phases were separated by centrifugation. Residual proteins and salt were removed by ethanol precipitation (3×). The final extracts were spotted on a thin-layer silica plate (Kieselgel 60F₂₅₄, Merck, Darmstadt, Germany) and developed with 94% (v/v) chloroform/methanol. Spots corresponding to 1-naphthol and 1,3-DHN were identified by comparison to the pure compounds and confirmed using an electrospray ionization-quadrupole ion trap mass spectrometer (LCQ, Finnigan, Bremen, Germany). No other hydroxylated naphthalenes were observed at significant concentrations in the product mixtures. Relative concentrations were determined from the integrated spot densities on the thin layer chromatogram (AlphaImager 2000).

Mutagenic PCR and StEP recombination. Mutagenic PCR was performed in a 100 μl reaction as described³¹, using 0.7 mM MnCl₂, 14 pmol of each primer, 5 units of *Taq* polymerase (Boehringer Mannheim), 0.01% gelatin, 20 pmol of template DNA (P450_{cam}-pCWori(+)). PCR was performed in a thermocycler (PTC200, MJ Research, Waltham, Massachusetts) for 30 cycles (94°C, 30 s; 45°C, 30 s; 72°C, 2 min). Sequences of forward and reverse primers are 5'-CATCGATGCTTAGG AGGTCATATG-3' and 5'-TCATGTTTGACAGCTTATCATCGAT-3'. The StEP method^{25,31} was used to recombine five mutant P450_{cam} genes from the first generation, under slightly mutagenic conditions (0.3 mM MnCl₂).

P450_{cam} purification and assay. Wild-type and M7-8H P450_{cam} were purified as described³². P450 content was estimated by measuring absorbance at 392 nm, with $\epsilon = 104 \text{ mM}^{-1} \text{ cm}^{-1}$. P450 fractions eluted from a Q-Sepharose (Pharmacia) column were pooled (average purity, $A_{392}/A_{280} = 1.4\text{--}1.6$) and washed three times with fresh 10 μM Tris buffer (pH 7.3). Assay conditions were: (200 μl total) 100 μM dibasic phosphate buffer (pH 9.0), 15% (v/v) pure ethanol, 30 μM naphthalene, 1.75 μM HRP (E.C.1.11.1.7, type II; Sigma), 100 μM H₂O₂, 0.031 μM (or 0.282 μg) purified P450. Fluorescence at 465 nm (excitation at 350 nm) was measured in a 96-well microfluorimeter. One unit is defined as the value of 1 a.u. fluorescence increase at 465 nm per min per μg P450.

Received 14 December 1998; accepted 11 May 1999.

- Nelson, D. R. *et al.* P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6**, 1–42 (1996).
- Cerniglia, C. E. Biodegradation of polycyclic aromatic hydrocarbons. *Biodegradation* **3**, 351–368 (1992).
- Cook, D. L. & Atkins, W. M. Enhanced detoxication due to distributive catalysis and toxic thresholds: a kinetic analysis. *Biochemistry* **36**, 10801–10806 (1997).
- Waxman, D. J. & Chang, T. K. H. In *Cytochrome P450: Structure, Mechanism and Biochemistry* 2nd edn (ed. Ortiz de Montellano, P. R.) 391–417 (Plenum, New York, 1995).
- Gonzalez, F. J. & Nebert, D. W. Evolution of the P450-gene superfamily—animal plant warfare, molecular drive and human genetic differences in drug oxidation. *Trends Genet.* **6**, 182–186 (1990).
- Sono, M., Roach, M. P., Coulter, E. D. & Dawson, J. H. Heme-containing oxygenases. *Chem. Rev.* **96**, 2841–2887 (1996).
- Jounaidi, Y., Hecht, J. E. D. & Waxman, D. J. Retroviral transfer of human cytochrome P450 genes for oxazaphosphorine-based cancer gene therapy. *Cancer Res.* **58**, 4391–4401 (1998).
- Wackett, L. P., Sadowsky, M. J., Newman, L. M., Hur, H.-G. & Li, S. Metabolism of polyhalogenated compounds by a genetically engineered bacterium. *Nature* **368**, 627–629 (1994).
- Faber, K. *Bioretransformations in Organic Chemistry* 208–220 (Springer, Berlin, 1997).

- Duport, C., Spagnoli, R., Degryse, E. & Pompon, D. Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. *Nature Biotechnol.* **16**, 186–189 (1998).
- Arnold, F. H. Design by directed evolution. *Acc. Chem. Res.* **31**, 125–131 (1998).
- Nordblom, G. D., White, R. E. & Coon, M. J. Studies on hydroperoxide-dependent substrate hydroxylation by purified liver microsomal cytochrome P450. *Arch. Biochem. Biophys.* **175**, 524–533 (1976).
- Hrycay, E. G., Gustafsson, J.-A., Ingelman-Sundberg, M. & Ernster, L. Sodium periodate, sodium chlorite, organic hydroperoxides and hydrogen peroxide as hydroxylating agents in steroid hydroxylation reactions catalyzed by partially purified cytochrome P450. *Biochem. Biophys. Res. Commun.* **66**, 209–216 (1975).
- Hummel, W. & Kula, M.-R. Dehydrogenases for the synthesis of chiral compounds. *Eur. J. Biochem.* **184**, 1–13 (1989).
- Seelbach, K. *et al.* A novel, efficient regenerating method of NADPH using a new formate dehydrogenase. *Tetrahedron Lett.* **37**, 1377–1380 (1997).
- Henriksen, A. *et al.* Structural interactions between horseradish peroxidase C and the substrate benzhydroxamic acid determined by X-ray crystallography. *Biochemistry* **37**, 8054–8060 (1998).
- Lin, Z., Thorsen, T. & Arnold, F. H. Functional expression of horseradish peroxidase in *Escherichia coli* by directed evolution. *Biotechnol. Prog.* **15**, 467–471 (1999).
- Gunsalus, I. C., Meeks, J. R., Lipscomb, J. D., Debrunner, P. G. & Munck, E. In *Molecular Mechanisms of Oxygen Activation* (ed. Hayaishi, O.) 559–613 (Academic, New York, 1974).
- England, P. A., Harford-Cross, C. F., Stevenson, J.-A., Rouch, D. A. & Wong, L.-L. The oxidation of naphthalene and pyrene by cytochrome P450_{cam}. *FEBS Lett.* **424**, 271–274 (1998).
- Poulos, T. L. & Raag, R. Cytochrome P450_{cam}: crystallography, oxygen activation, and electron transfer. *FASEB J.* **6**, 674–679 (1992).
- Graham-Lorence, S. & Peterson, J. A. P450s: Structural similarities and functional differences. *FASEB J.* **10**, 206–214 (1996).
- Denu, J. M. & Tanner, K. G. Specific and reversible inactivation of protein tyrosine phosphatases by hydrogen peroxide: evidence for a sulfenic acid intermediate and implications for redox regulation. *Biochemistry* **37**, 5633–5642 (1998).
- van Deuren, M. P. J., van Rantwijk, F. & Sheldon, R. A. Selective oxidations catalyzed by peroxidases. *Tetrahedron* **53**, 13183–13220 (1997).
- Cramer, A., Raillard, S.-A., Bermudez, E. & Stemmer, W. P. C. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288–291 (1998).
- Zhao, H., Giver, L., Shao, Z., Alfholter, J. A. & Arnold, F. H. Molecular evolution by staggered extension process (StEP) *in vitro* recombination. *Nature Biotechnol.* **16**, 258–261 (1998).
- Moore, J. C. & Arnold, F. H. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nature Biotechnol.* **14**, 458–467 (1996).
- Giver, L., Gershenson, A., Freskgard, P. O. & Arnold, F. H. Directed evolution of a thermostable esterase. *Proc. Natl Acad. Sci. USA* **95**, 12809–12813 (1998).
- Zhao, H. & Arnold, F. H. Directed evolution converts subtilisin E into a functional equivalent of the thermolysin. *Protein Eng.* **12**, 47–53 (1999).
- McKay, C. P. & Hartman, H. Hydrogen peroxide and the evolution of oxygenic photosynthesis. *Origins Life Evol. Biosphere* **21**, 157–163 (1991).
- Samuilov, V. D. Photosynthetic oxygen: the role of H₂O₂. A review. *Biochemistry (Moscow)* **62**, 451–454 (1997).
- Zhao, H., Moore, J. C., Volkov, A. A. & Arnold, F. H. In *Manual of Industrial Microbiology and Biotechnology* 2nd edn (eds Demain, A. L. & Davies, J. E.) 597–604 (ASM Press, Washington DC, 1999).
- Unger, B. P., Gunsalus, I. C. & Sligar, S. G. Nucleotide-sequence of the *Pseudomonas-putida* cytochrome P-450_{cam} gene and its expression in *Escherichia coli*. *J. Biol. Chem.* **261**, 1158–1163 (1986).

Acknowledgements. We thank J. H. Richards for encouragement and support, P. R. Ortiz de Montellano for discussions, and G. Bandara for assistance with purification. This work was supported by the Biotechnology Research and Development Corporation and by the Office of Naval Research. H.J. received partial support from the Korea Research Foundation.

Correspondence and requests for materials should be addressed to F.H.A. (e-mail: frances@chem.ealtech.edu).

Reassessment of ice-age cooling of the tropical ocean and atmosphere

S. W. Hostetler* & A. C. Mix†

* US Geological Survey, 200 SW 35th Street, Corvallis, Oregon 97331-5503, USA

† College of Oceanic & Atmospheric Science, Ocean Administration Building 104, Oregon State University, Corvallis, Oregon 97331-5503, USA

The CLIMAP¹ project's reconstruction of past sea surface temperature inferred limited ice-age cooling in the tropical oceans. This conclusion has been controversial, however, because of the greater cooling indicated by other terrestrial and ocean proxy data^{2–6}. A new faunal sea surface temperature reconstruction, calibrated using the variation of foraminiferal species through time, better represents ice-age faunal assemblages and so reveals greater cooling than CLIMAP in the equatorial current systems of the eastern Pacific and tropical Atlantic oceans⁷. Here we explore the climatic implications of this revised sea surface temperature field for the Last Glacial Maximum using an atmospheric general circulation model. Relative to model results obtained using

CLIMAP sea surface temperatures, the cooler equatorial oceans modify seasonal air temperatures by 1–2 °C or more across parts of South America, Africa and southeast Asia and cause attendant changes in regional moisture patterns. In our simulation of the Last Glacial Maximum, the Amazon lowlands, for example, are cooler and drier, whereas the Andean highlands are cooler and wetter than the control simulation. Our results may help to resolve some of the apparent disagreements between oceanic and continental proxy climate data. Moreover, they suggest a wind-related mechanism for enhancing the export of water vapour from the Atlantic to the Indo-Pacific oceans, which may link variations in deep-water production and high-latitude climate changes to equatorial sea surface temperatures.

Modern foraminiferal assemblages in the tropics are poor analogues for those of the ice-age ocean in the eastern equatorial Pacific and Atlantic oceans. Such limited representation of the true ice-age fauna in these areas decreases the sensitivity of traditional faunal transfer-functions (which use statistical relationships to predict temperature from species distributions) in these regions, and leads to underestimates of sea surface temperature (SST) changes (Fig. 1a). A new reconstruction (referred to as the 'OSU' reconstruction)⁷ resolves the analogue problem by calibrating modern (core-top) faunal assemblages on the basis of ancient samples. In this reconstructed SST anomaly field (Last Glacial Maximum, LGM, minus modern; Fig. 1b), significant cooling spans the Atlantic Ocean from the eastern boundary regions along Africa (4–5 °C cooling) to the Caribbean (~3 °C cooling). In the Pacific Ocean, cooling of up to 5 °C extends westward from Peru along the Equator, roughly similar to a perpetual La Niña climate state. Near the centre of the subtropical gyres, the OSU reconstruction indicates only a small SST cooling (or even slight warming) relative to present, roughly similar to CLIMAP¹.

The largest changes in the LGM fauna near the Equator involve species associated with the eastern boundary currents, suggesting enhanced circulation of the subtropical gyres during the LGM⁷. This inferred circulation is consistent both with data suggesting a cooler and better-ventilated thermocline⁸, and with some coupled atmosphere–ocean climate models that display enhanced poleward oceanic heat transport into the subtropics⁹.

We evaluate the effects of this new LGM reconstruction on tropical climate in a series of sensitivity tests with an atmospheric general circulation model, AGCM (GENESIS 2.01, atmospheric resolution 3.75° by 3.75°, surface resolution 2° by 2° latitude and longitude¹⁰). Three simulations were conducted: (1) modern climate with prescribed SSTs (the control simulation), (2) 21,000

years ago (21 kyr ago), with prescribed CLIMAP SSTs (the CLIMAP simulation), and (3) 21 kyr ago with the new tropical SST reconstruction⁷ substituted for CLIMAP in the area 20° N–20° S, 140° W–20° E (the OSU simulation). All simulations used modern vegetation¹¹ and either modern or LGM sea ice and ice sheets^{12,13}, orbital parameters, and atmospheric CO₂ concentration (300 parts per million by volume, p.p.m.v., modern, 200 p.p.m.v. LGM). We did not change the CLIMAP temperatures in the western Pacific or Indian oceans because revisions to SSTs in these areas would probably be much smaller than the changes we propose for the eastern Pacific and Atlantic¹⁴.

The cooler equatorial SSTs in the OSU simulation induce widespread atmospheric anomalies relative to the CLIMAP simulation (Fig. 2). Cooling is evident in summer (June, July, August; JJA) and winter (December, January, February; DJF) over the central Andes, lowland South America, tropical and subtropical Africa, eastern Africa, the Arabian peninsula, and southern Asia. In the OSU simulation, a large positive sea-level pressure anomaly associated with cooler SSTs spans the Equator in the tropical Atlantic and eastern Pacific oceans. The change in wind flow induced by this higher pressure, and by opposing circulation patterns in the northern and southern hemispheres (clockwise and anticlockwise out of high-pressure centres, respectively), is manifested in enhanced westward flow over northern South America into the Pacific (JJA) and by enhanced eastward flow in the mid-latitudes over the South Atlantic (DJF and JJA). Elsewhere, modified pressure gradients change surface winds over Africa, the Arabian peninsula, and the Indian Ocean. The monsoonal winds over India and southeast Asia are weaker in the OSU simulation than in the control but stronger than those of the CLIMAP simulation; they are thus more consistent with an observed small reduction in Arabian Sea upwelling during the LGM¹⁵.

The equatorial Walker circulation is characterized as a zonal pattern of ascending air over Indonesia, South America, and Africa and descending air over the eastern Pacific, Atlantic and western Indian oceans. Vertical motions are estimated here based on changes in the vertical pressure gradient in the model. Relative to CLIMAP, in the OSU simulation rising motions are weakened over South America and strengthened over Indonesia, while sinking motions are enhanced over the eastern Pacific, Atlantic, and western Indian oceans. Such changes, together with modified winds at the surface and aloft, indicate a stronger Walker circulation in the OSU simulation, which contributes to changes in the distribution of atmospheric moisture (Fig. 2).

In both seasons, net moisture (precipitation minus evaporation, P – E) anomalies are negative (that is, drier) in the OSU simulation relative to CLIMAP near and just downwind of the cooler SSTs (Fig. 2). Seasonal drying in the OSU simulation extends well inland over the Amazon basin and Central America. Drying also extends eastward over central Africa to the Indian Ocean. Off the Equator, the thermal gradients focus moisture from the relatively warm subtropical gyres along the edges of the cool anomalies, resulting in net moisture increases over the Andes and northern Mexico, generally consistent with terrestrial data¹⁶. During boreal summer, in the OSU simulation wetter conditions also exist in the monsoon regions of India and southeast Asia, as a result of pressure-related changes in wind fields and stronger rising and sinking motions in the Indo-Pacific sector of the Walker circulation, however, both the OSU and CLIMAP simulations produce summer drying relative to the control simulation.

Net export of atmospheric water vapour from the Atlantic drainage helps to maintain modern global thermohaline ocean circulation. For the control simulation, our estimate of net freshwater export from the Atlantic to the Pacific and Indian oceans (integrated vertically from 1,000 to 200 hPa) is ~0.1 Sv (1 Sv = 10⁶ m³ s⁻¹), a value similar to other model estimates but lower than values inferred from atmospheric data¹⁷. In the model,

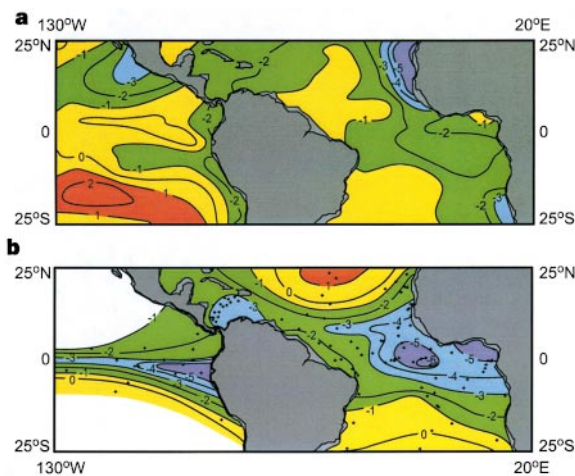
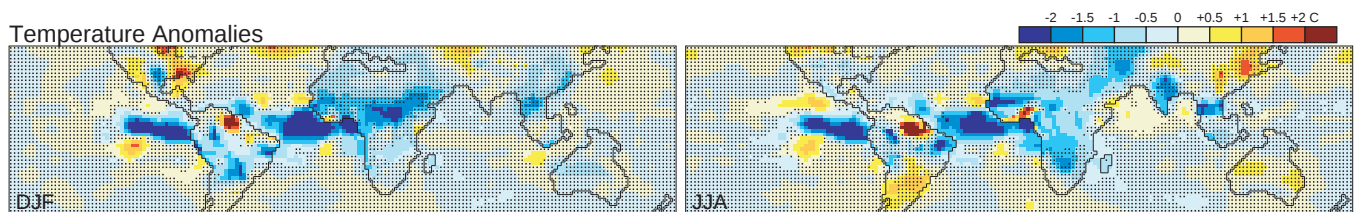
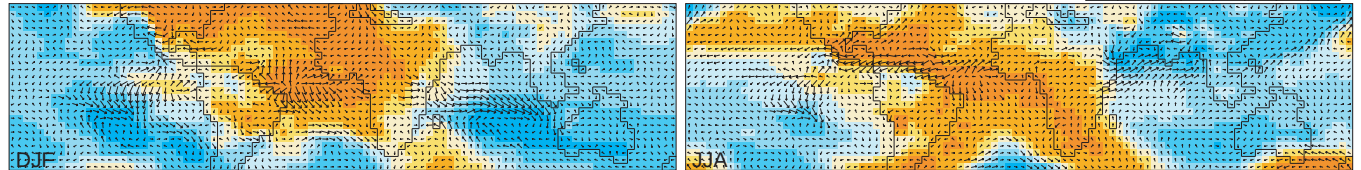


Figure 1 Sea surface temperature anomalies estimated from foraminiferal species. **a**, CLIMAP LGM¹ minus modern; **b**, OSU LGM⁷ minus modern. Points indicate core locations of LGM samples. White areas indicate no data.

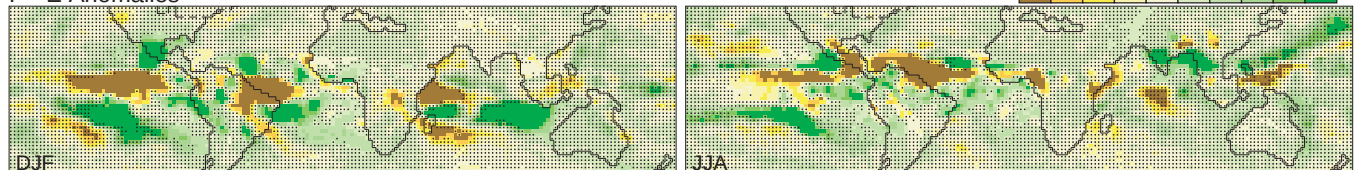
Temperature Anomalies



SLP and Surface Wind Anomalies



P - E Anomalies



Meridionally Averaged Specific Humidity Anomalies (35N - 35S)

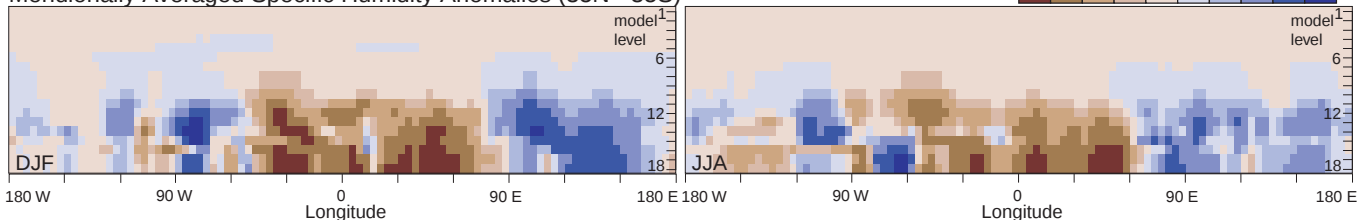


Figure 2 Simulated LGM climate anomalies for the region between 40° N and 40° S latitude. Data are OSU LGM minus CLIMAP LGM, averages of years 4–15 of the model runs. DJF, December–February; JJA, June–August. Anomalies in stippled areas are not statistically different from the model's interannual variability (at the $p = 0.9$ level)³⁰. Temperature anomalies are for surface (2-m) air tempera-

tures; SLP is sea-level pressure; surface wind anomalies are vector differences from the lowest model level; P – E is precipitation minus evaporation at the surface. Specific humidity anomalies are a vertical cross-section of the atmosphere, averaged meridionally from 35° N to 35° S; equally spaced model levels expand the view of the lower troposphere. Level 11 corresponds approximately to 500 hPa.

freshwater export doubles to ~ 0.2 Sv in the CLIMAP simulation, and quadruples to ~ 0.4 Sv in the OSU simulation. The enhanced export in the latter simulation is attributable in part to stronger westward winds across northern South America, which compensate for suppressed evaporation from a cooler Atlantic. Although some of the moisture transported from the Atlantic augments precipitation over South America, there is a net loss of fresh water from the Atlantic basin. Over time, the loss would tend to increase the salinity

of the Atlantic and thereby influence North Atlantic Deep Water production, suggesting a mechanism for linking climate oscillations in low and high latitudes¹⁸.

Estimates of mean annual air temperatures from analysis of rare gases in tropical ground water suggest LGM cooling of 4–6 °C in lowland Brazil, Nigeria and Vietnam^{3–4}. The OSU simulation agrees with the groundwater estimates in Nigeria and Vietnam, but achieves less cooling than the groundwater data in lowland Brazil

Table 1 Simulated climate anomalies at the Last Glacial Maximum

Location	OSU – CLIMAP ΔT (°C)		OSU – Control ΔT (°C)		Data	OSU – CLIMAP $\Delta(P - E)$ (mm d ⁻¹)		OSU – Control $\Delta(P - E)$ (mm d ⁻¹)	
	DJF	JJA	DJF	JJA		DJF	JJA	DJF	JJA
Andes region	-1	-1	-4	-5		+1	0	+3	0
Huascarán	-1	0	-5	-5	-0.9 km ELA ²	+3	-1	+5	-2
Titicaca	-1	-1	-4	-4	Cooler, wetter ²⁰	+1	+1	+3	+1
Amazon region	0	+1	-2	-1		-5	0	3	-2
Lowland Brazil	0	-2	-2	-3	-5 °C (ref. 3)	-6	0	-4	-1
Sahel region	-1	-1	-8	-2		0	0	0	-1
Nigeria	-1	-1	-7	-3	-4 to -5 °C (ref. 4)	0	-3	0	-2
Kilimanjaro	0	0	-2	-4	-0.9 km ELA ²¹	0	-1	-1	+3
SE Asia region	-1	-1	-4	-2		0	+1	+2	-7
Vietnam	-1	-2	-5	-3	-4 to -5 °C (ref. 4)	0	+2	+2	-6
Texas	-2	0	-7	-6	-5 °C (ref. 4)	+1	0	+1	+1
Georgia	+1	0	-4	-5	-4 °C (ref. 4)	0	0	+2	-1
New Guinea	0	0	-2	-3	-0.9 km ELA ²²	0	+1	+1	-1
Hawaii	0	0	+1	+1	-0.9 km ELA ²³	0	-1	-2	+3

Surface temperature anomalies (ΔT) and precipitation minus evaporation anomalies ($\Delta(P - E)$) among the model runs are computed as averages over 2° latitude by 2° longitude land grid points (except Hawaii, which is all ocean grid points) within the specified regions. For the more specific locations, the anomalies are averaged over grid boxes within 2° latitude and longitude. The regions and locations are defined as: Andes 7° S to 25° S, 77° W to 65° W; Huascarán 9° S, 77° W; Titicaca 17° S, 69° W; Amazon 3° N to 5° S, 55° W to 45° W; lowland Brazil 7° S, 41° W; Sahel 21° N to 9° N, 5° W to 17° E; Nigeria 9° N, 11° E; Kilimanjaro 3° S, 37° E; SE Asia 25° N to 9° N, 89° E to 109° E; Vietnam 11° N, 105° E; Texas 29° N, 99° W; Georgia 33° N, 83° W; New Guinea 5° S, 141° E; Hawaii 19° N, 155° W. DJF, December–February; JJA, June–August; ELA, equilibrium-line altitude.

(Table 1). In the Amazon lowlands the OSU simulation is substantially drier than the CLIMAP simulation (up to 5 mm d^{-1} in DJF), and is consistent with evidence for Pleistocene aridity¹⁹. Both the CLIMAP and OSU simulations are in agreement with groundwater temperature estimates from Texas and Georgia⁴.

Glacier data from the Andes²⁰, Kilimanjaro²¹, New Guinea²², and Hawaii²³ indicate that equilibrium-line altitudes were $\sim 900 \text{ m}$ lower than at present ($\sim 780 \text{ m}$ reduction relative to lower sea level). But limiting dates for these advances range widely (20–15 kyr ago for Huascarán, 30–10 kyr for Kilimanjaro, >15 kyr for New Guinea and 22–9 kyr for Hawaii), so the synchronicity of lowered equilibrium-line altitudes across the tropics remains uncertain. For Kilimanjaro and Huascarán, the OSU simulation is $3\text{--}5^\circ\text{C}$ cooler than the control, yielding approximate temperature-related depressions of equilibrium-line altitudes of $550\text{--}900 \text{ m}$ (based on a nominal tropical lapse rate of $5.5^\circ\text{C km}^{-1}$). In both areas net moisture at the LGM is greater than that of the control, which would further enhance glacier growth. The OSU simulation is thus consistent with regional glacier advances in the Andes and east Africa. We did not modify CLIMAP SSTs near New Guinea and Hawaii, and as a result temperatures in the OSU simulation are essentially the same as those of the CLIMAP simulation (3°C cooler and 1°C warmer than the control, respectively) and perhaps inconsistent with local glaciation, indicating the need for further study of these regions²⁴.

The OSU simulation helps to resolve some, but not all, disagreements between land and ocean data. We did not consider climate feedbacks associated with LGM vegetation, and these may yield further modelled cooling over land²⁵. Although the ice-age tropics in the OSU simulation are generally cooler and drier than the control, some regions such as the Andean highlands are cooler and wetter, suggesting substantial variation within the tropics both regionally and with altitude, consistent with recent data compilations²⁶. The LGM cooling in the OSU simulation agrees well with recent ocean and atmosphere–ocean model simulations over the eastern Pacific, the equatorial Atlantic and regions of Africa and Asia^{9,27–29}. To the extent that our AGCM results are model-independent, this agreement suggests convergence of data and models in these regions. In the western Pacific, the OSU reconstruction agrees with one ocean simulation²⁹ and one coupled atmosphere–ocean simulation²⁸, but disagrees with a coupled atmosphere–ocean simulation that yields a cooling of $4\text{--}6^\circ\text{C}$ (ref. 27). Nevertheless, our results highlight the potentially widespread influences of regional SST changes in circulation patterns and moisture fluxes associated with tropical–extratropical temperature gradients, and understore the need to reconstruct both the amplitude and the geographical distribution of LGM climate changes in much greater detail. □

Received 28 January; accepted 5 May 1999.

- CLIMAP Project Members. Seasonal reconstruction of the earth's surface at the last glacial maximum. (Map & Chart Ser. MC-3, Geol. Soc. Am., Boulder, Colorado, 1981).
- Thompson, L. G. *et al.* Late glacial stage and Holocene tropical ice core records from Huascarán, Peru. *Science* **269**, 46–50 (1995).
- Stute, M. *et al.* Cooling of tropical Brazil (5°C) during the last glacial maximum. *Science* **269**, 379–383 (1995).
- Stute, M. *et al.* Glacial paleotemperature records for the tropics derived from noble gases dissolved in groundwater (abstr). *Eos* **78**, F44 (1997).
- Guilderson, T. P., Fairbanks, R. G. & Rubenstone, J. L. Tropical temperature variations since 20,000 years ago: modulating interhemispheric climate change. *Science* **263**, 663–665 (1994).
- Rind, D. & Peteet, D. Terrestrial conditions at the last glacial maximum and CLIMAP sea-surface temperature estimates: are they consistent? *Quat. Res.* **24**, 1–22 (1985).
- Mix, A. C., Morey, A., Pisias, N. G. & Hostetler, S. Foraminiferal faunal estimates of paleotemperature: Circumventing the no-analog problem yields cool ice age tropics. *Paleoceanography* **14**, 350–359 (1999).
- Slowey, N. & Curry, W. B. Glacial-interglacial differences in circulation and carbon cycling within the upper western North Atlantic. *Paleoceanography* **10**, 715–732 (1995).
- Ganopolski, A., Rahmstorf, S., Petoukhov, V. & Claussen, M. Simulation of modern and glacial climates with a coupled global model of intermediate complexity. *Nature* **391**, 351–356 (1998).
- Thompson, S. L. & Pollard, D. A global climate model (GENESIS) with a land-surface-transfer scheme (LSX), 1, present-day climate. *J. Clim.* **8**, 732–761 (1995).
- Dorman, J. L. & Sellers, P. J. A global climatology of albedo, roughness length, and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB). *J. Appl. Meteorol.* **28**, 833–855 (1989).
- Clark, P. U., Licciardi, J. M., MacAyeal, D. R. & Jenson, J. W. Numerical reconstruction of a soft-bedded Laurentide ice sheet during the last glacial maximum. *Geology* **24**, 679–682 (1996).

- Peltier, W. R. Ice age paleotopography. *Science* **265**, 195–201 (1994).
- Thunell, R. C., Anderson, D. M., Gellar, D. & Miao, Q. Sea-surface temperature estimates for the tropical western Pacific during the last glaciation and their implications for the Pacific warm pool. *Quat. Res.* **41**, 255–264 (1994).
- Prell, W. L. & Kutzbach, J. E. Monsoon variability over the past 150,000 years. *J. Geophys. Res.* **92**, 8411–8425 (1987).
- Markgraf, V. in *Global Climates Since the Last Glacial Maximum* (eds Wright, H. E. *et al.*) 357–385 (Univ. Minnesota Press, Minneapolis, 1993).
- Zaucker, F. & Broecker, W. S. The influence of atmospheric moisture transport on the fresh water balance of the Atlantic drainage basin: General circulation model simulations and observations. *J. Geophys. Res.* **97**, 2765–2773 (1992).
- Zaucker, F., Stocker, T. F. & Broecker, W. S. Atmospheric freshwater fluxes and their effect on the global thermohaline circulation. *J. Geophys. Res.* **99**, 12443–12457 (1994).
- Servant, M. *et al.* Tropical forest changes during the late Quaternary in African and South American lowlands. *Glob. Planet. Change* **7**, 25–40 (1993).
- Klein, A. G., Seltzer, G. O. & Isacks, B. L. Modern and last local glacial maximum snowlines in the Central Andes of Peru, Bolivia, and Northern Chile. *Quat. Sci. Rev.* **18**, 63–84 (1999).
- Downie, C. & Wilkinson, P. *The Geology of Kilimanjaro* (Univ. Sheffield Press, 1972).
- Löffler, E. Pleistocene glaciation in Papua and New Guinea. *Z. Geomorphol.* **13**, 32–58 (1972).
- Porter, S. C. Chronology of Hawaiian glaciations. *Science* **195**, 61–63 (1977).
- Lee, K. E. & Slowey, N. C. Cool surface waters of the subtropical North Pacific Ocean during the last glacial. *Nature* **397**, 512–514 (1999).
- Crowley, T. J. & Baum, S. K. Effect of vegetation on an ice-age climate model simulation. *J. Geophys. Res.* **102**, 16463–16480 (1997).
- Farrera, I. *et al.* Tropical climates at the last glacial maximum: A new synthesis of terrestrial palaeoclimate data, I, Vegetation, lake levels, and geochemistry. *Clim. Dyn.* (in the press).
- Bush, A. B. G. & Philander, S. G. H. The role of ocean-atmosphere interactions in tropical cooling during the last glacial maximum. *Science* **279**, 1341–1344 (1998).
- Weaver, A. J., Eby, M., Fanning, A. F. & Wiebe, E. C. Simulated influence of carbon dioxide, orbital forcing, and ice sheets on the climate of the last glacial maximum. *Nature* **394**, 847–853 (1998).
- Bigg, G. R., Wadley, M. R., Stevens, D. P. & Johnson, J. A. Simulation of two last glacial maximum ocean states. *Paleoceanography* **13**, 340–351 (1998).
- Chervin, R. M. Interannual variability and seasonal predictability. *J. Atmos. Sci.* **43**, 233–241 (1986).

Acknowledgements. We thank P. Bartlein for discussions and the design of Fig. 2; P. Clark, N. Pisias, D. Pollard for discussions; P. Valdez for comments and suggestions; and A. Morey and D. Zahnle for assistance. Model simulations were conducted at the National Center for Atmospheric Research, the Environmental Protection Agency, and the College of Oceanic and Atmospheric Sciences at Oregon State University. The US Geological Survey (S.W.H.) and The National Science Foundation (A.C.M.) supported this research.

Correspondence and requests for materials should be addressed to S.W.H. (e-mail: steve@ucar.edu).

Gold concentrations of magmatic brines and the metal budget of porphyry copper deposits

T. Ulrich*, D. Günther*† & C. A. Heinrich*

* Isotope Geology and Mineral Resources, Department of Earth Sciences, ETH Zentrum 8092 Zürich, Switzerland

Porphyry copper–molybdenum–gold deposits are the most important metal resources formed by hydrothermal processes associated with magmatism. It remains controversial, however, whether the metal content of porphyry-style and other magmatic–hydrothermal deposits is dominantly controlled by metal partitioning between magma and an exsolving magmatic fluid phase^{1,2} or by scavenging of metals from solid upper-crustal rocks by surface-derived fluids³. It also remains unknown to what degree the metal content in such deposits is affected by selective mineral precipitation from the ore fluid. Extremely saline fluids⁴, precipitating quartz and ore minerals in veins have been inferred to have a significant magma-derived component, on the basis of geological⁵, isotopic^{6,7} and experimental evidence^{8,9}. Here we report gold and copper concentrations of single fluid inclusions in quartz, determined by laser-ablation inductively coupled plasma mass spectrometry. The results show that the Au/Cu ratio of primary high-temperature brines is identical to the bulk Au/Cu ratio in two of the world's largest copper–gold ore bodies. This indicates that the bulk metal budget of such deposits is primarily controlled by the composition of the incoming fluid, which is, in turn, likely to be controlled by the crystallization

† Present address: Laboratory of Inorganic Chemistry, ETH Zentrum 8092 Zürich, Switzerland.

process in an underlying magma chamber.

Porphyry-type deposits occur along certain segments of the circumpacific and Alpine–Himalayan belts and are hosted in a dense network of quartz–sulphide veins cutting subvolcanic stocks of calcalkaline composition. Copper and gold are the most strongly enriched elements in these deposits, but to a highly variable degree. Copper concentrations in the ore bodies are up to about 1 weight per cent (typical enrichment factor ~ 200 compared to average crustal abundance), whereas gold concentrations vary from less than 0.05 p.p.m. in gold-poor deposits to greater than 1 p.p.m. in the economically most attractive porphyry copper–gold deposits (enrichment factors ~ 20 – 400)¹⁰. Gold and copper distributions within individual deposits are closely correlated at all scales, ranging from an intimate textural intergrowth at the mineral grain scale (Fig. 1) to highly correlated element concentrations with characteristic Au/Cu ratios in ore samples across the deposit (Fig. 2c). This close chemical and mineralogical association of Au and Cu in each deposit, with vastly different absolute concentrations and significant variations of Cu/Au ratios among different deposits, makes the two ore metals the obvious and most direct tracers for the hydrothermal enrichment process—provided that their concentration in the hydrothermal fluid can be determined independently.

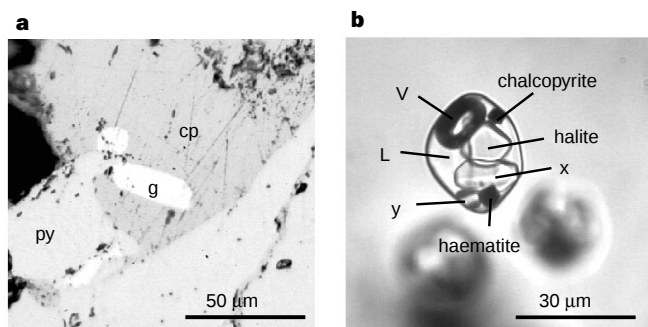


Figure 1 Photomicrographs of a polished section of ore and of a polyphase brine inclusion. **a**, The intimate association of gold (g) and chalcopyrite (cp) ± pyrite (py) in porphyry-related ore deposits is shown by small gold grains occurring within chalcopyrite, possibly exsolved from a high-temperature solid solution. **b**, A typical high-temperature, brine inclusion, which upon cooling and partial loss of H_2 had exsolved into aqueous liquid (L), a vapour bubble (V), a salt crystal, small chalcopyrite and haematite grains and several unidentified crystals (x, y; sample from Bajo de la Alumbrera, Argentina).

Here we report quantitative data on the concentration of gold in porphyry-mineralizing fluids. For this pilot study, we examined the highest-temperature, most saline brine inclusions from Grasberg (the world's richest porphyry copper–gold deposit located in Irian Jaya¹¹, Indonesia), and from Bajo de la Alumbrera (another giant gold-rich copper deposit in Argentina¹², Fig. 1). Quartz veins in pervasively altered porphyries containing hydrothermal quartz + K-feldspar + magnetite + minor chalcopyrite ($CuFeS_2$) with associated gold were selected from the deepest part in the centre of the hydrothermal systems; we wanted to sample inclusions that represent the magmatic–hydrothermal fluids before substantial modification by fluid/rock reactions or mixing with upper-crustal waters. Coexisting low-density inclusions give evidence for the simultaneous entrapment of a high-temperature vapour phase of low salinity. Vapour and brine phases were trapped in the same quartz crystals as isolated inclusions, and have been studied individually by *in situ* trace-element analysis using laser ablation inductively coupled plasma mass spectrometry (ICP-MS) and microthermometry.

The measurements were performed on regular-shaped fluid inclusions of 10–45 μm size in pre-ore vein quartz. Polyphase brine inclusions from Bajo de la Alumbrera (Fig. 1b), from several inclusion assemblages within the same quartz vein, have similar salinities of 58–65 wt% NaCl (equivalent) and homogenization temperatures of 550 to 650 °C. Brine inclusions along several trails in the Grasberg sample have apparent salinities of 68–76 wt% NaCl (equivalent) and homogenization temperatures above 600 °C. The salinity of vapour-rich inclusions from Grasberg (7.4 ± 3 wt% NaCl equivalent; high compared with experimental phase relations due to minor co-entrapment of brine) was determined from the temperature of final ice melting. No microthermometry analyses could be performed on the low-density vapour-rich inclusions from Alumbrera owing to their small liquid content.

For laser ablation ICP-MS analysis, selected individual inclusions were 'drilled' out of the polished quartz samples with an ArF excimer laser^{13,14}. To detect Au we used a miniaturized helium transport system that carries the ablated inclusion content into a modified quadrupole ICP-MS for multi-element data acquisition¹⁵. Helium increases the efficiency of aerosol transport from the ablation spot into the plasma, and enhances the signal-to-background ratio for gold by an order of magnitude. Plasma and interface conditions of the ICP-MS were specifically adjusted to give maximum response on ¹⁹⁷Au. A reduced set of elements (silicon, sodium, copper, gold and arsenic), and an increased

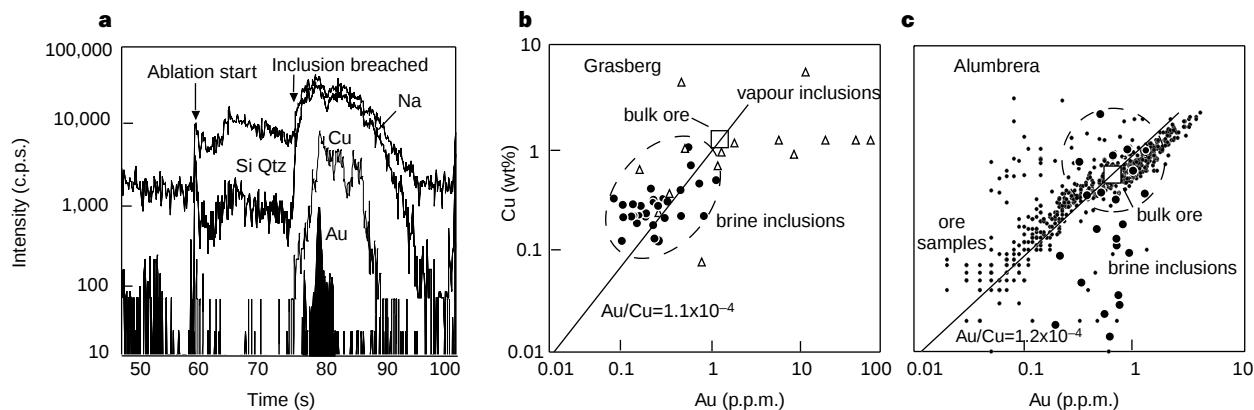


Figure 2 Transient laser ablation ICP-MS signal and Au versus Cu plots from Grasberg and Bajo de la Alumbrera. Shown are results from early, high-temperature brine inclusions (filled circles surrounded by dashed line) and vapour-rich inclusions (open triangles); also shown for comparison are the metal concentration ratio of the bulk ore bodies (open square with diagonal line) and representative ore samples (small points). **a**, Time-resolved ICP-MS signal showing the progress of laser ablation through host quartz, followed by the

breaching of a brine inclusion and complete ablation, including an internally precipitated copper sulphide crystal to which gold is attached. Consistent element concentrations result from integration of these signals, as shown by results obtained from the giant porphyry Cu–Au deposits of Grasberg (**b**) and Alumbrera (**c**). Both deposits contain high-temperature brine inclusions with Au and Cu in the same concentration ratio as the bulk ore.

dwelt time on ^{197}Au (50 ms, compared to 10 ms for the other elements) further improved the detection limit of gold. Limits of detection as low as 0.1 p.p.m. Au in 25- μm inclusions were determined, corresponding to a total mass of $\sim 10^{-15}$ g Au.

Transient signals record the entire ablation process of each inclusion, and demonstrate unambiguously that Au and Cu, and other trace, minor and major elements (including Na) originated from within the fluid inclusion (Fig. 2a). Consistently short and overlapping peaks for Cu and Au show that gold is attached to a tiny chalcopyrite (CuFeS_2) crystal, precipitated inside the inclusions during cooling of the host mineral to room temperature after trapping of a homogeneous fluid sample at high temperature and pressure (Fig. 1b). Absolute trace-element concentrations are obtained by signal integration, calibration with an external standard, and internal standardization of absolute concentrations using Na, which was determined previously for each inclusion by microthermometry¹⁴. Uncertainties of individual analyses are estimated to be $\pm 30\%$, increasing for small inclusions ($< 10 \mu\text{m}$) or concentrations near the detection limit. Table 1 gives average concentrations of Au, Cu, Na, and As in brine and vapour inclusions, together with additional elements analysed in the same inclusion assemblages.

Au versus Cu concentrations of brine inclusions are shown in Fig. 2b and c (filled circles), together with the bulk Au and Cu contents of the two giant ore bodies (open squares) and a spatially representative suite of ore-sample assays illustrating the variable but highly correlated Cu and Au grades within the Bajo de la Alumbrera deposit (Fig. 2c, small dots). The data from Grasberg (Fig. 2b) show that the average Au/Cu ratio in brine inclusions (0.9×10^{-4} by weight) is almost identical to that of the bulk deposit (1.1×10^{-4} ; ref. 16). Similarly, at Alumbrera (Fig. 2c), one group of relatively Cu-rich brine inclusions overlaps the bulk metal ratio of the deposit (1.2×10^{-4} ; refs 17, 18), although some petrographically identical inclusions have lower Cu concentrations at similar levels of Au of ~ 0.6 p.p.m. Vapour inclusions coexisting with brine inclusions in the Grasberg sample (Fig. 2b, triangles) have comparable Au/Cu ratios but contain, on average, about ten times higher concentrations of both metals¹⁹. Associated (but not necessarily coexisting) vapour inclusions in the Alumbrera sample have high copper concentrations but no detectable gold. We do not know whether the compositional variation among the brine inclusions at Alumbrera is due to preferential partitioning of Cu into the vapour phase, or to partial precipitation as a Cu–Fe sulphide. Aware of this uncertainty, we interpret the brines with the ore-matching metal ratio (filled circles marked by dashed lines in Fig. 2b and c) as the best approximation to the initial composition of the main metal-introducing fluid in both hydrothermal systems.

Table 1 Average concentrations of fluid inclusions

Element		Alumbrera		Grasberg	
		Brine	Vapour*	Brine	Vapour
Au	p.p.m.	0.79 ± 0.39	< 0.53	0.26 ± 0.18	10.17 ± 6.20
Cu	wt%	0.76 ± 0.49	3.30 ± 1.20	0.30 ± 1.00	1.20 ± 1.30
Na	wt%	16.00 ± 2.00	1.70 ± 0.14	16.00 ± 1.00	3.00 ± 1.00
K	wt%	12.50 ± 2.70	0.72 ± 0.24	15.40 ± 4.70	1.30 ± 0.20
Mn	wt%	1.50 ± 0.30	0.14 ± 0.06	2.40 ± 0.80	0.20 ± 0.10
Fe	wt%	8.50 ± 1.90	1.30 ± 0.15	13.00 ± 3.50	1.00 ± 0.50
Zn	wt%	1.40 ± 0.30	0.12 ± 0.03	1.30 ± 0.40	0.15 ± 0.07
Pb	wt%	0.45 ± 0.13	0.02 ± 0.01	0.50 ± 0.25	0.04 ± 0.01
Rb	p.p.m.	750 ± 175	25 ± 6	960 ± 240	80 ± 30
Sr	p.p.m.	85 ± 30	10 ± 7	640 ± 350	30 ± 20
Mo	p.p.m.	70 ± 60	< 300	600 ± 120	60 ± 20
Ag	p.p.m.	45 ± 70	< 40	$1,200 \pm 300$	100 ± 40
Cs	p.p.m.	60 ± 15	20 ± 20	70 ± 35	5 ± 1
As	p.p.m.	20 ± 20	n.a.	20 ± 10	190 ± 220
Ba	p.p.m.	95 ± 25	40 ± 20	380 ± 170	10 ± 10

* Estimated from experimental phase relations and microthermometric data of brine inclusions; absolute element concentrations of these vapour inclusions are less certain than concentration ratios. n.a., not analysed.

The correspondence of Au/Cu ratios in the primary ore fluids and in the bulk deposits is remarkable, considering that individual concentrations of the two metals in common crustal fluids are likely to vary independently over many orders of magnitude. In conjunction with the variable but highly correlated ore grades, our observation leads to the conclusion that copper and gold must have been transported together in the same fluid, and were co-precipitated almost quantitatively within the volumes of the economic ore bodies. This strongly indicates that the composition of the incoming magmatic brine exerted the dominant chemical control on the final bulk composition of the two deposits. Co-transportation of Cu and Au is consistent with experimental data indicating that both metals in their +1 valence state form stable chloride complexes in high-temperature saline solutions^{9,20}. Their co-precipitation is probably driven by fluid cooling²¹, and is possibly aided by initial incorporation of Au into Cu–Fe sulphide solid solutions²².

While co-precipitation of Cu and Au in both deposits was very efficient, metal precipitation was also highly selective, as shown by the high concentrations of other metals in the brine. Concentrations of lead and particularly zinc in the brines are comparable to, or even higher than, the Cu concentration, and yet neither lead nor zinc is significantly enriched in the ore bodies. This demonstrates that very large quantities of ore metals are flushed through ore deposits, and ultimately become dispersed unless steep thermal and chemical gradients provide a driving force for efficient mineral precipitation.

The source processes for Cu and Au can be constrained by first-order mass-balance considerations, based on the average concentration of the bulk-ore-matching fluid (0.79 p.p.m. Au and 0.76 wt% Cu) and the total metal content of the Alumbrera deposit (460 tonnes (t) of Au, and 3.7×10^6 t of Cu)^{17,18}. To advect this mass of metals to the deposit, at least 5.8×10^8 t of brine are required. Calcalkaline melts are estimated to contain ~ 2 p.p.b. Au or less^{23,24}. To supply the gold content of the deposit would therefore require at least 2.3×10^{11} t of magma. This is much larger than the quantity of porphyry that hosts and immediately underlies the ore body, and requires the presence of a deeper-seated magma chamber of 100 km³ volume or more, consistent with tentative aeromagnetic evidence and exposed remnants of a large stratovolcano²⁵. To source the copper from the same magma chamber would only require extraction of 15 p.p.m., that is, one-quarter of the typical copper content of andesitic magmas (60 p.p.m.; ref. 26). Comparing the 5.8×10^8 t of brine used for metal transport with 2.3×10^{11} t of magma required to source the gold, we obtain a brine/magma ratio of $1/400 = 0.2\%$; this is a small fraction of the water content (~ 4 wt%) estimated for typical porphyry copper source magmas based on petrological constraints^{27,28}. Considering experimental fluid/melt partitioning data^{27,29–32}, we speculate that the highly Cl-, Cu- and Au-enriched ore brine could represent the first fraction of fluid that exsolved from magma that contained initially ~ 4 wt% H₂O and had a normal H₂O/Cl ratio of about ten (ref. 27). The magma was emplaced and started to crystallize at conditions near fluid saturation, about 3 km below the current exposure level of the deposit. Alternatively, the highly metal-charged brine might indicate that an unusually gold- and/or chlorine-rich (H₂O/Cl $\ll 10$) magma was involved in the formation of the two porphyry copper–gold deposits that we studied. □

Received 30 December 1998; accepted 13 April 1999.

1. Candela, P. A. *Rev. Econ. Geol.* **4**, 203–221 (1989).
2. Cline, J. S. & Bodnar, R. J. Can economic porphyry copper mineralization be generated by a typical calc-alkaline melt? *J. Geophys. Res.* **96**, 8113–8126 (1991).
3. Sheets, R. W., Nesbitt, B. E. & Muehlenbachs, K. Meteoric water component in magmatic fluids from porphyry copper mineralization, Babine Lake area, British Columbia. *Geology* **24**, 1091–1094 (1996).
4. Roedder, E. *Fluid Inclusions* (Mineralogical Soc. of America, 1984).
5. Hedenquist, J. W. & Lowenstern, J. B. The role of magmas in the formation of hydrothermal ore deposits. *Nature* **370**, 519–527 (1994).
6. Sheppard, S. M. F., Nielsen, R. L. & Taylor, H. P. J. Hydrogen and oxygen isotope ratios in minerals from porphyry copper deposits. *Econ. Geol.* **66**, 515–542 (1971).
7. Sheppard, S. M. F., Nielsen, R. L. & Taylor, H. P. J. Oxygen and hydrogen ratios of clay minerals from porphyry copper deposits. *Econ. Geol.* **64**, 755–777 (1969).

8. Hemley, J. J., Cygan, G. L., Fein, J. B., Robinson, G. R. & D'Angelo, W. M. Hydrothermal ore-forming processes in the light of studies in rock-buffered systems: I. Iron-copper-zinc-lead sulfide solubility relations. *Econ. Geol.* **87**, 1–22 (1992).
9. Seward, T. M. & Barnes, H. L. in *Geochemistry of Hydrothermal Ore Deposits* 3rd edn (ed. Barnes, H. L.) 435–486 (Wiley, New York, 1997).
10. Taylor, S. R. & McLennan, S. M. *The Continental Crust: its Composition and Evolution* (Blackwell Scientific, Oxford, 1985).
11. McDonald, G. D. & Arnold, L. C. Geological and geochemical zoning of the Grasberg igneous complex, Irian Jaya, Indonesia. *J. Geochem. Explor.* **50**, 143–178 (1994).
12. Ulrich, T. in *Actas X Congreso Latinoamericano de Geología and VI Congreso Nacional de Geología Económica* 239 (Universidad de Buenos Aires, 1998).
13. Günther, D., Frischknecht, R., Heinrich, C. A. & Kahlert, H.-J. Capabilities of an argon fluoride 193nm excimer laser for laser ablation inductively coupled plasma mass spectrometry microanalysis of geological materials. *J. Anal. Atom. Spectrosc.* **12**, 939–944 (1997).
14. Günther, D., Audétat, A., Frischknecht, R. & Heinrich, C. A. Quantitative analysis of major, minor and trace elements in fluid inclusions using laser ablation-inductively coupled plasma-mass spectrometry (LA-ICP-MS). *J. Anal. Atom. Spectrosc.* **13**, 263–270 (1998).
15. Günther, D. & Heinrich, C. A. Enhanced sensitivity in laser ablation ICP-mass-spectrometry using helium-argon mixtures as aerosol carrier. *J. Anal. Atom. Spectrosc.* (submitted).
16. Meinert, L. D., Hefton, K. K., Mayes, D. & Tasiran, I. Geology, zonation, and fluid evolution of the big gossan Cu-Au skarn, Ertsberg district, Irian Jaya. *Econ. Geol.* **92**, 509–534 (1997).
17. Wall, V. J. Bajo de la Alumbrera (Argentina). A world class copper gold deposit. *Rev. Asoc. Argentina Geol. Econ.* **11**, 92–93 (1997).
18. Forrester, P. MIM Holdings Limited, financial and production results. (cited Nov. 98) (<http://www.mimholdings.com.au>) (1998).
19. Heinrich, C. A., Günther, D., Audétat, A., Ulrich, T. & Frischknecht, R. Metal fraction between magmatic brine and vapor, and the link between porphyry-style and epithermal Cu-Au deposits. *Geology* (in the press).
20. Hemley, J. J. & Hunt, J. P. Hydrothermal ore-forming processes in the light of studies in rock-buffered systems: II. Some general geologic applications. *Econ. Geol.* **87**, 23–43 (1992).
21. Gammons, C. H. & Williams-Jones, A. E. Chemical mobility of gold in the porphyry-epithermal environment. *Econ. Geol.* **92**, 45–59 (1997).
22. Simon, G. Partitioning of gold between bornite and chalcopyrite in porphyry-copper-gold deposits: A thermodynamic and experimental approach. *Econ. Geol.* (submitted).
23. Connors, K. A., Noble, D. C., Bussey, S. D. & Weiss, S. I. Initial gold contents of silicic volcanic rocks: Bearing on the behavior of gold in magmatic systems. *Geology* **21**, 937–940 (1993).
24. Wallace, D. A. The precious-metals potential of the Rocky volcanics in the Lachlan fold belt. *Austral. Geol. Surv. Org. Res. Newsl.* **24**, 17–18 (1996).
25. Sasso, A. M. & Clark, A. H. The Farallón Negro group, northwestern Argentina: Magmatic, hydrothermal and tectonic evolution and implications for Cu-Au metallogeny in the Andean back-arc. *Soc. Econ. Geol. Newsl.* **34**, 1–18 (1998).
26. Gill, J. *Orogenic Andesites and Plate Tectonics* (Springer, New York, 1981).
27. Cline, J. S. in *Porphyry Copper Deposits of the American Cordillera* (eds Pierce, F. W. & Bolm, J. G.) 69–82 (Arizona Geological Soc., Tucson, 1995).
28. Naney, M. T. Phase equilibria of rock-forming ferromagnesian silicates in granitic systems. *Am. J. Sci.* **283**, 993–1033 (1983).
29. Candela, P. A. & Piccoli, P. M. in *Magmas, Fluids, and Ore Deposits* (ed. Thompson, J. F. H.) 101–127 (Mineralogical Soc. Canada, Victoria, British Columbia, 1995).
30. Kilinc, I. A. & Burnham, C. W. Partitioning of chloride between a silicate melt and coexisting aqueous phase from 2 to 8 kilobars. *Econ. Geol.* **67**, 231–235 (1972).
31. Nakano, T. & Urabe, T. Calculated composition of fluids released from a crystallizing granitic melt: Importance of pressure on the genesis of ore forming fluid. *Geochem. J.* **23**, 307–319 (1989).
32. Shinohara, H. Exsolution of immiscible vapor and liquid phases from a crystallizing silicate melt: Implications for chlorine and metal transport. *Geochim. Cosmochim. Acta* **58**, 5215–5221 (1994).

Acknowledgements. We thank MIM Exploration for financial and logistic support at Bajo de la Alumbrera, and Freeport Ltd for guidance at Grasberg. We thank H. Barnes, J. Hedenquist, S. Kesler and S. Matthai for critically reading the manuscript. Project and equipment funding by ETH Zürich and Schweizerischer Nationalfond is acknowledged.

Correspondence and requests for materials should be addressed to C.H. (e-mail: heinrich@erdw.ethz.ch).

A diapsid skull in a new species of the primitive bird *Confuciusornis*

Lianhai Hou*, Larry D. Martin†, Zhonghe Zhou*†, Alan Feduccia‡ & Fucheng Zhang*

* Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, P.O. Box 643, Beijing 100044, China
 † Natural History Museum and the Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas 66045, USA
 ‡ Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280, USA

Since the description of *Confuciusornis* (the oldest beaked bird) in 1995, based on three partial specimens, large numbers of complete skeletons have been recovered^{1,2}. Most new material of *Confuciusornis*^{3,4} can be assigned to a single sexually dimorphic species, *C. sanctus*. Here we report a new species based on a

remarkably well preserved skeleton with feathers and, for the first time in the Mesozoic record, direct evidence of the shape of a horny beak. It has a complete and large preserved postorbital that has a broad contact with the jugal bone. This character is presently only known in *Confuciusornis*, and may confirm previous suggestions of a postorbital in *Archaeopteryx*⁵. The squamosal is in tight contact with the postorbital. These two bones form an arch dividing the upper and lower temporal fenestrae, as in other diapsid reptiles⁶. The presence of a typical diapsid cheek region with two openings in *Confuciusornis* may preclude the presence of prokinesis (upper jaw mobility against the braincase and orbital area), a feeding adaptation found in most modern birds. The presence of a horny beak, characteristic of modern birds, coupled with a primitive temporal region provides new evidence for a mosaic pattern in the early evolution of birds.

Aves Linnaeus 1758

Sauriurae Haeckel 1866

Confuciusornithiformes Hou *et al.* 1995

Confuciusornithidae Hou *et al.* 1995

Confuciusornis Hou *et al.* 1995

Confuciusornis dui sp. nov.

Etymology. The species name is dedicated to Mr. Wenya Du, who collected and donated the specimen to the Institute of Vertebrate Paleontology and Paleoanthropology (IVPP) for scientific research. **Holotype.** A nearly complete skeleton. IVPP Collection Number V 11553.

Paratype. IVPP 11521, a partial skeleton consisting of a sternum, ribs, vertebrae, pelvis, femora and tail.

Horizon and locality. A two-metre thick interval within the Yixian Formation (Late Jurassic-Early Cretaceous); Libalanggou, Zhangjiying, Beipiao, Liaoning, northeast China.

Diagnosis. The holotype, a presumed male, is about 15% smaller than the holotype of *C. sanctus* (a small individual and presumed female). Large male individuals of *C. sanctus* are about 30% larger than the new species. The mandible is more slender anteriorly,



Figure 1 Cast of the elongate tail feathers of *C. dui* (IVPP Collection V 11553) whitened with ammonium chloride, showing that it was probably male.

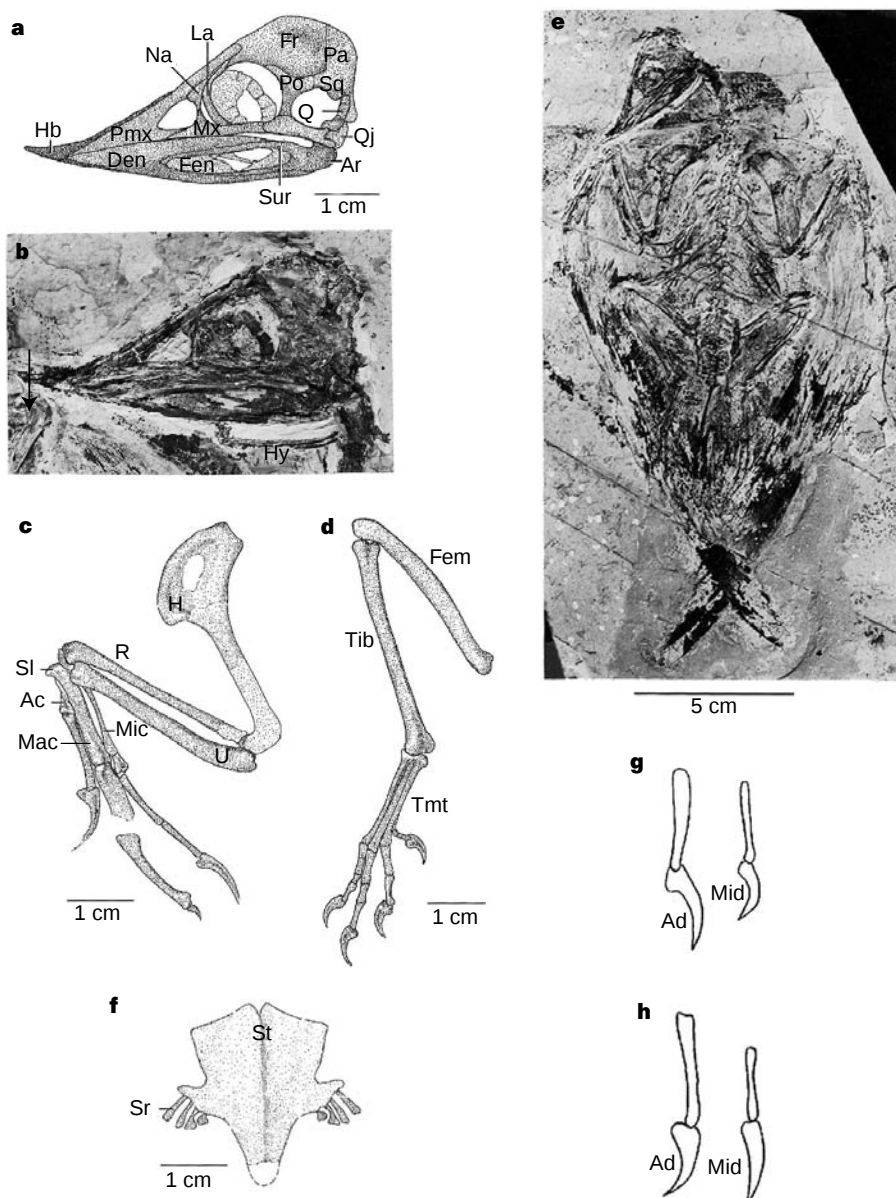


Figure 2 *Confuciusornis dui*. **a**, Reconstruction of the skull. **b**, Photograph of the skull (IVPP Collection V 11553); the arrow indicates the horny beak. **c**, Forelimb (IVPP Collection V 11553). **d**, Hindlimb. **e**, Photograph of the holotype (IVPP Collection V 11553). **f**, Reconstruction of the sternum and sternal ribs (based on IVPP Collection V 11521). **g**, Alular and minor manual digits of *C. sanctus*. **h**, Alular and minor manual digits of *C. dui*. Abbreviations: Ac, alular metacarpal; Ad, alular manual digit; Ar,

articular; Den, dentary; Fen, fenestra of the lower jaw; Fem, femur; Fr, frontal; H, humerus; Hb, horny beak; Hy, hyoid bone; J, jugal; La, lacrimal; Mac, major metacarpal; Mic, minor metacarpal; Mid, distal minor manual digit; Mx, maxilla; Na, nasal; Pa, parietal; Po, postorbital; Q, quadrate; Qj, quadratojugal; R, radius; Sl, semilunate bone (fused with major metacarpal); Sq, squamosal; Sr, sternal ribs; St, sternum; Sur, surangular; Tib, tibiotarsus; Tmt, tarsometatarsus; U, ulna.

without the distinctive anteroventral expansion of the dentary found in *C. sanctus*. The upper jaw is also more pointed anteriorly than that of *C. sanctus*. The claw of the alular digit is not enlarged as in *C. sanctus*. The sternum is more elongate with an anterior notch and a pair of short lateral processes. The tarsometatarsus is relatively shorter than in *C. sanctus*, and is shorter than the pygostyle.

Measurements of holotype (mm). Wing chord, 201; length of lower jaw 40; length of humerus, 42; length of ulna, 39; length of femur, 35; length of tibiotarsus, 41; length of tarsometatarsus, 19.5; length of pygostyle, 23; length of carpometacarpus, 19.

Description. This is a smaller, more gracile species than *C. sanctus*. The tarsometatarsus is fused proximally, and the semilunate bone is fused to the major metacarpal, indicating that this is an adult specimen. Like other *Confuciusornis*, it lacks teeth. Impressions of a

single pair of long tail feathers (Fig. 1) indicate that IVPP V11553 was a male.

The skull (Fig. 2a, b) is well preserved and shows some critical features of *Confuciusornis*. The impression of the horny rhamphotheca is preserved (Fig. 2a, b). The horny bill extends in front of the bony core with a long and pointed tip. The distal end of the beak curves dorsally and the bill was not raptorial. It is likely that *Confuciusornis* was an herbivore. The nasal processes of the premaxillae separate the nasals and overlap the anterior end of the frontals. The lacrimal joins the nasal above a small slit-shaped antorbital fenestra. The postorbital is large and 'Y' shaped; it contacts the jugal, squamosal and frontal. The contact on the jugal for the postorbital appears to be a rounded process. The quadrate is similar to that of *Archaeopteryx* and lacks an orbital

process. The quadratojugal seems to be small and 'L' shaped, as in *Archaeopteryx*. The maxillary process of the premaxilla tapers posteriorly and overlaps the maxilla below the middle of the narial opening. The maxilla has a dorsal ascending process in contact with the nasal, forming the anterior margin of the antorbital fenestra. There is a 'V'-shaped groove formed by backwards-extending processes of the dentary surrounding a large fenestra in the surangular.

The neck is relatively short. The exact number of cervical vertebrae in the new species is unclear, but it is probably fewer than eight. The vertebrae are short with deep pleurocoels. The pygostyle is longer than the tarsometatarsus and is well fused. The exact number of vertebrae that are fused into the pygostyle is unknown in the new species, but we have counted between 8 and 10 in *C. sanctus*. The sternal plate, which is visible on the holotype but best seen in the referred specimen (Fig. 2f), is longer than it is wide. It is relatively flat and lacks an obvious keel, as in *C. sanctus*. There are a pair of short lateral processes that serve as attachment sites for four short sternal ribs (Fig. 2f). Gastralria are present. The flight feathers (Fig. 2e) are very asymmetric with a wing chord of about 201 mm. The twin tail feathers are clearly visible in low-incidence lighting and resemble those of other *Confuciusornis* (Fig. 1).

The humerus is typical of *Confuciusornis* in having a very expanded proximal end with an oval depression (it is not clear whether any of the many specimens of *Confuciusornis* shown with a humeral perforation had that character before preparation). The humerus (Fig. 2c) is slightly longer than the ulna, as in *Archaeopteryx* but different from later birds where the ulna is usually longer than the humerus. The ulna is nearly twice as wide as the radius. The semilunate bone is fused to the major metacarpal, forming a carpometacarpus, but not to the alular metacarpal, indicating that the semilunate bone in *Confuciusornis*, as in modern birds, is a single distal carpal. The major and minor metacarpals are of nearly equal length and the major metacarpal is more robust than the other two. The minor metacarpal is significantly narrower proximally than distally, as in *C. sanctus*. There are three clawed fingers and the claw of the alular digit is about the same size as that of the minor digit, rather than about one-third larger as in *C. sanctus* (Fig. 2g, h). The middle manual claw is reduced, as in other *Confuciusornis*, although it is not completely preserved. The first phalanx of the alular digit is almost as long as the major and minor metacarpals. As is typical of *Confuciusornis*, the first phalanx of the minor digit is short (the first and second are short in *Archaeopteryx*); the others of the same digit are long, and the second is longer than the penultimate phalanx.

The pelvis is narrow and opisthopubic. The femur (Fig. 2d) is almost as long as the tibiotarsus. The tarsometatarsus is fused only proximally, as in *C. sanctus* and other saururine birds. The foot has highly recurved claws and a reversed hallux.

Discussion. The evolution of the skull in early birds is less known than that of the postcranial skeleton. This is particularly true of the temporal region. The cheek region in modern birds is reduced, owing to an expanded brain and enlarged orbit. Fossils of the oldest bird, *Archaeopteryx*, have unfortunately not preserved the post-orbital and provide little information about the transition from the typical diapsid skull to the situation in modern birds.

Confuciusornis dui preserves a large 'Y'-shaped postorbital with a broad contact with the jugal bone. The postorbital process of the jugal does not lie behind the postorbital in contrast to that of *Dromaeosaurus*⁷. The postorbital has not been found in *Archaeopteryx*, but its presence has been extrapolated on the basis of the postorbital process on the jugal bone and the shape of the squamosal^{5,8}. The jugal bone in *Confuciusornis* is strong and relatively short compared with that of modern birds, where it is slender and fused with the quadratojugal bone, forming a rod-like bar. The postorbital was probably absent in the Chinese Early

Cretaceous enantiornithine *Cathayornis*^{9–11}. The only other Mesozoic bird that has retained this structure is an Early Cretaceous bird from Spain¹². However, in this bird the postorbital is much more reduced and has lost its contact with the jugal bar.

The squamosal is in tight contact with the postorbital. They form an intact intertemporal arch separating the upper and lower temporal fenestrae. The unreduced postorbital and its contacts with the jugal bone and the squamosal might prevent prokinesis; this may be confirmed by the fact that the quadrate has no orbital process and the quadratojugal is reduced so it cannot participate in a prokinetic push-rod. The presence of two derived quadrate systems indicates that a streptostylic quadrate with an orbital process would not have been a feature of the first bird.

Although *Confuciusornis* retains an intact diapsid temporal region, it shows several cranial modifications also found in more advanced birds. The enlargement of the frontal and reduction of the parietal make the frontal the major contributor to the skull internal to the upper arch, whereas in typical diapsid animals the parietal extends further anteriorly. The dorsal process of the jugal bone is low and has a rounded contact with the postorbital. The quadrate of *Confuciusornis* is single-headed; it lacks the prominent orbital process that is present in modern birds but is also lacking in the Early Cretaceous enantiornithine *Cathayornis*¹¹.

Although the sternum has no obvious keel, the shape of the wing and flight feathers are highly adapted for powered flight. *Confuciusornis* probably lived in large aggregations along the margin of a freshwater lake, as shown by hundreds of individuals found in a limited area. There is abundant evidence for lush forested conditions and the freshwater mudstones contain a variety of conchostracans, insects, fishes, amphibians, reptiles and mammals. The long tails on the presumed males seem to preclude much activity on the ground or the water surface and, along with the highly recurved pedal claws and reversed hallux, indicate that *Confuciusornis* was a perching bird. None of the many specimens preserve stomach contents but it seems likely that *Confuciusornis* was herbivorous.

The Early Cretaceous saururine birds from younger geological sections are smaller than the Late Jurassic *Archaeopteryx* and the Late Jurassic–Early Cretaceous *Confuciusornis*. The new species of *Confuciusornis* is smaller than *C. sanctus*. It has been proposed that size reduction was important in the early evolution of avian flight¹³. The early ornithurine birds and Early Cretaceous enantiornithines are smaller than *Archaeopteryx* and *Confuciusornis*. In addition, the new species has a more elongated sternum.

The findings of different species of *Confuciusornis* show that the oldest known beaked bird was not only abundant but had also diversified, as had *Archaeopteryx*, with at least two species in both genera⁵. A recently found enantiornithine bird¹⁴ from the same locality as *Confuciusornis* and *Liaoningornis* indicates that Enantiornithes had a longer history than previously known. Thus, the origin and early diversification of birds might be earlier and more significant than expected^{15,16}.

All known Late Jurassic–Early Cretaceous ornithurines such as *Chaoyangia* and enantiornithines have retained teeth in the jaw¹, and only *Confuciusornis* lost its teeth completely in favour of a horny beak. It provides an unusual mosaic of primitive and derived features with the anterior skull more like that of a modern bird than in any other Late Jurassic–Early Cretaceous bird, but with a temporal region nearly unchanged from its remote archosaurian ancestors. Like the early evolution of mammals, the original diversification of birds was probably also a complicated bush with many extinct lines that may at one time have been more advanced in some features than their ultimately more successful contemporaries¹⁷. The combination of distinctively advanced and primitive features found in the skull provides new evidence for a mosaic pattern in the early evolution of birds. *Confuciusornis* is not the progenitor of either modern birds or later enantiornithines,

but must be regarded as an early twig in a bush-like radiation of birds. □

Received 15 February; accepted 21 April 1999.

- Hou, L., Martin, L. D., Zhou, Z. & Feduccia, A. Early adaptive radiation of birds: evidence from fossils from northeastern China. *Science* **274**, 1164–1167 (1996).
- Peters, D. S. Ein nahezu vollständiges Skelette eines urtümlichen Vogels aus China. *Natur und Museum* **126**, 298–302 (1996).
- Hou, L., Zhou, Z., Gu, Y. & Zhang, H. *Confuciusornis sanctus*, a new Late Jurassic sauriurine bird from China. *Chin. Sci. Bull.* **40**, 1545–1551 (1995).
- Hou, L., Zhou, Z., Martin, L. D. & Feduccia, A. A beaked bird from the Jurassic of China. *Nature* **377**, 616–618 (1995).
- Elzanowski, A. & Wellnhofer, P. Cranial morphology of *Archaeopteryx*: evidence from the seventh skeleton. *J. Vert. Paleontol.* **16**, 81–94 (1996).
- Reisz, R. R. A diapsid reptile from the Pennsylvanian of Kansas. *Special Publ. Nat. Hist. Mus. Univ. Kansas* **7**, 1–74 (1981).
- Colbert, E. H. & Russell, D. A. The small Cretaceous dinosaur *Dromaeosaurus*. *Amer. Mus. Novit.* **2380**, 1–49 (1969).
- Wellnhofer, P. Das fünfte Skelettexemplar von *Archaeopteryx*. *Palaeontographica A* **147**, 169–216 (1974).
- Zhou, Z., Jin, F. & Zhang, J. Preliminary report on a Mesozoic bird from Liaoning, China. *Chin. Sci. Bull.* **37**, 1365–1368 (1992).
- Zhou, Z. The discovery of Early Cretaceous birds in China. *Cour. Forsch. Inst. Senckenb.* **181**, 9–22 (1995).
- Martin, L. D. & Zhou, Z. *Archaeopteryx*-like skull in enantiornithine bird. *Nature* **389**, 556 (1997).
- Sanz, J. L. et al. An Early Cretaceous bird from Spain and its implication for the evolution of avian flight. *Science* **276**, 1543–1546 (1997).
- Zhou, Z. & Hou, L. *Confuciusornis* and the early evolution of birds. *Vertebr. Palasiat* **36**, 136–146 (1998).
- Hou, L., Martin, L. D., Zhou, Z. & Feduccia, A. *Archaeopteryx* to opposite birds—missing link from the Mesozoic of China. *Vertebr. Palasiat* **37** (in the press).
- Feduccia, A. *The Origin and Evolution of Birds* (Yale Univ. Press, New Haven, 1996).
- Chatterjee, S. *The Rise of Birds* (John Hopkins University Press, Baltimore, 1997).
- Martin, L. D. & Miao, D. in *Short Papers of the Sixth Symposium on Mesozoic Terrestrial Ecosystems and Biota* 217–219 (China Ocean Press, Beijing, 1995).

Acknowledgements. We thank L. Witmer and S. Chatterjee for critical and helpful comments and reviews, and D. Miao for assistance and suggestions. The Chinese Natural Science Foundation, the National Geographic Society (U.S.) and the Grand Project of the Chinese Academy of Sciences supported fieldwork in Liaoning Province, northeast China. M. Tanner did the drawings and J. Chorn the photographs. D. Miao and J. Chorn critically read the manuscript.

Correspondence and requests for materials should be addressed to L.D.M. (e-mail: ldmartin@falcon.cc.ukans.edu).

Cultures in chimpanzees

A. Whiten*, J. Goodall†, W. C. McGrew‡, T. Nishida§, V. Reynolds||, Y. Sugiyama¶, C. E. G. Tutin#*, R. W. Wrangham** & C. Boesch††

* Scottish Primate Research Group, School of Psychology, University of St Andrews, St Andrews KY16 9JU, UK

† Gombe Stream Research Centre, P.O. Box 185, Kigoma, Tanzania

‡ Department of Zoology and Department of Sociology, Gerontology and Anthropology, Miami University, Oxford, Ohio 45056, USA

§ Laboratory of Human Evolution Studies, Kyoto University, Kyoto 606-01, Japan

|| Institute of Biological Anthropology, Oxford University, 58 Banbury Road, Oxford OX2 6QS, UK

¶ Primate Research Institute, Kyoto University, Inuyama 484-8506, Japan

Centre Internationale de Recherche Médicales de Franceville,

BP 769 Franceville, Gabon

* Department of Biological Sciences, University of Stirling, Stirling FK9 4LA, UK

** Department of Anthropology, Harvard University, Cambridge, Massachusetts 02138, USA

†† Max-Planck Institute for Evolutionary Anthropology, Inselstrasse 22, 04301 Leipzig, Germany

As an increasing number of field studies of chimpanzees (*Pan troglodytes*) have achieved long-term status across Africa, differences in the behavioural repertoires described have become apparent that suggest there is significant cultural variation^{1–7}. Here we present a systematic synthesis of this information from the seven most long-term studies, which together have accumulated 151 years of chimpanzee observation. This comprehensive analysis reveals patterns of variation that are far more extensive than have previously been documented for any animal species except humans^{8–11}. We find that 39 different behaviour patterns, including tool usage, grooming and courtship behaviours, are

customary or habitual in some communities but are absent in others where ecological explanations have been discounted. Among mammalian and avian species, cultural variation has previously been identified only for single behaviour patterns, such as the local dialects of song-birds^{12,13}. The extensive, multiple variations now documented for chimpanzees are thus without parallel. Moreover, the combined repertoire of these behaviour patterns in each chimpanzee community is itself highly distinctive, a phenomenon characteristic of human cultures¹⁴ but previously unrecognised in non-human species.

Culture is defined in very different ways in different academic disciplines¹⁵. At one extreme, some cultural anthropologists insist on linguistic mediation, so that culture is constrained to be a uniquely human phenomenon¹⁶. In the biological sciences, a more inclusive definition is accepted, in which the significance of cultural transmission is recognized as one of only two important processes that can generate evolutionary change: inter-generation transmission of behaviour may occur either genetically or through social learning, with processes of variation and selection shaping biological evolution in the first case and cultural evolution in the second. From this perspective, a cultural behaviour is one that is transmitted repeatedly through social or observational learning to become a population-level characteristic¹⁷. By this definition, cultural differences (often known as 'traditions' in ethology) are well established phenomena in the animal kingdom and are maintained through a variety of social transmission mechanisms¹⁸. Well documented examples include dialects in song-birds^{12,13}, sweet-potato washing by Japanese macaques (*Macaca fuscata*) at Koshima¹⁹, and stone handling by Japanese macaques at Arashiyama²⁰. However, each case refers to variation in only a single behaviour pattern.

Tabulations of population differences amongst chimpanzees have indicated that multiple behavioural variants may exist^{2–7}. However, these tabulations have been based on published reports, which, although they record the presence of behaviours, remain problematic in three respects: they are incomplete; they frequently do not clarify the extent to which each behaviour pattern is habitual in the community; and they do not systematically document the absence of behaviour patterns present elsewhere. We therefore adopted a different strategy in our attempt to provide a definitive assessment of what is now known of chimpanzee cultural variation.

Phase 1 of the study established a comprehensive list of candidate cultural variants, which are behaviours suspected by research workers to be specific to particular chimpanzee populations. Beginning with a list drawn from literature review by A.W. and C.B., the research directors of the major chimpanzee field projects (Table 1) added and defined unpublished candidate patterns. The patterns were then split and lumped as appropriate. This complex, collaborative and iterative process produced a listing of candidate cultural variants that were fully and consensually defined (see Supplementary Information; Table 1 gives abridged descriptions). The scope of this list, differentiating 65 categories of behaviour, represents a unique record of the inventiveness of wild chimpanzees.

In phase 2, the research directors assigned to each of these behaviour categories one of the following six codes, as applicable at their site: (1) customary, for which the behaviour occurs in all or most able-bodied members of at least one age-sex class (such as adult males); (2) habitual, for which the behaviour is not customary but has occurred repeatedly in several individuals, consistent with some degree of social transmission; (3) present, for which the behaviour is neither customary nor habitual but is clearly identified; (4) absent, for which the behaviour has not been recorded and no ecological explanation is apparent; (5) ecological explanation, for which absence is explicable because of a local ecological feature; and (6) unknown, for which the behaviour has not been recorded, but this may be due to inadequacy of relevant observational opportunities. These codings were cross-checked and confirmed by senior colleagues at each site. Our results are for the seven chimpanzee

groups with the most long-term observation record, so the 'unknown' code was seldom applicable (Table 1). These studies bring together a total of 151 years of direct observation (range 8–38 years), so our data summarize the enormous increase in our knowledge of chimpanzee behaviour achieved in the latter half of this century.

For any row in Table 1, the profile of codings of particular interest with respect to cultural variation is that in which behaviours are recorded as customary or habitual in some communities, yet absent at others. Three other classes of profile need to be recognized and discriminated from this.

First, seven behaviours proposed as potential cultural variants in

Table 1 Variation in occurrence of behaviour patterns across long-term study sites

			Site						
			Bs	Ta	Go	Ma	Mk	Kib	Bd
A	1	Investigatory probe (probe and sniff)	H	C	C	H	H	+	(-)
	2	Play start (invite play holding stem in mouth)	+	H	C	C	C	C	H
	3	Drag branch (drag large branch in display)	H	C	C	C	C	H	H
	4	Leaf-sponge (leaf mass used as sponge)	C	C	C	+	e	C	C
	5	Branch-clasp (clasp branch above, groom)	H	C	C	C	C	C	C
	6	Branch-shake (to attract attention, court)	C	C	C	C	C	H	C
	7	Buttress-beat (drum on buttress of tree)	C	C	C	C	C	C	C
B	8	Nasal probe (clear nasal passage with stick)	-	-	-	+	-	-	-
	9	Comb (stem used to comb through hair)	-	-	-	-	-	-	+
	10	Insect-pound (probe used to mash insect)	+	-	-	-	-	-	-
	11	Resin-pound (extract resin by pounding)	+	-	-	e?	e?	-	-
	12	Branch-hook (branch used to hook branch)	+	-	-	-	-	-	-
	13	Perforate (stout stick perforates termite nest)	-	e	-	-	-	e	e?
	14	Dig (stick used as spade to dig termite nest)	+	e	-	-	-	e	e?
	15	Brush-stick (probing stick with brush end)	-	-	-	-	-	-	-
	16	Seat-stick (stick protection from thorns)	-	-	e	e?	e?	e	e
	17	Stepping-stick (walking on sticks over thorns)	-	-	e	e?	e?	e	e
	18	Container (object used as container)	-	-	+	-	-	-	-
	19	Leaf-mop (leaves used to mop up insects)	-	-	+	-	+	e	e?
	20	Leaf-wipe (food wiped from skull etc.)	e?	+	+	-	-	-	-
	21	Leaf-brush (leaf used to brush away bees)	-	-	+	-	-	-	-
	22	Open and probe (perforate, then probe)	-	-	-	-	-	-	-
	23	Sponge push-pull (stick and sponge tool)	+	+	+	+	e	e	-
C	24	Algae-scoop (scoop algae using wand)	C	e	e	e	e	e	e
	25	Ground-night-nest (night-nests on ground)	(-)	e?	+	e?	e?	e?	+
	26	Anvil-prop (rock used to level anvil)	H	e	e	e	e	e	e
D	27	Food-pound onto wood (smash food)	C	C	C	-	-	e?	H
	28	Food-pound onto other (such as stone)	-	H	C	-	-	e?	-
	29	Nut-hammer, wood hammer on wood anvil	-	C	-	e	e	e?	e
	30	Nut-hammer, wood hammer on stone anvil	-	C	-	-	-	e?	e
	31	Nut-hammer, stone hammer on wood anvil	+	C	-	e	e	e?	e
	32	Nut-hammer, stone hammer on stone anvil	C	C	-	-	-	e?	e
	33	Nut-hammer, other (such as on ground)	-	H	-	-	-	e?	e
	34	Pestle-pound (mash palm crown with petiole)	C	-	-	e?	e?	e?	e?
	35	Club (strike forcefully with stick)	+	H	H	+	-	+	-
	36	Termite-fish using leaf midrib	+	e	-	-	C	e	e?
	37	Termite-fish using non-leaf materials	-	e	C	-	C	e	e?
	38	Ant-fish (probe used to extract ants)	+	-	+	C	C	-	-
	39	Ant-dip-wipe (manually wipe ants off wand)	+	-	C	-	-	-	-
	40	Ant-dip-single (mouth ants off stick)	C	C	+	-	-	-	-
	41	Fluid-dip (use of probe to extract fluids)	-	C	C	H	H	H	-
	42	Bee-probe (disable bees, flick with probe)	-	C	-	-	+	-	-
	43	Marrow-pick (pick bone marrow out)	-	C	-	-	-	-	-
	44	Lever open (stick used to enlarge entrance)	-	H	C	-	-	-	-
	45	Expel/stir (stick expels or stirs insects)	-	C	H	H	H	-	-
	46	Seat-vegetation (large leaves as seat)	+	H	-	-	-	+	-
	47	Fly-whisk (leafy stick used to fan flies)	-	H	+	-	-	-	H
	48	Self-tickle (tickle self using objects)	-	-	H	-	-	-	-
	49	Aimed-throw (throw object directionally)	C	C	C	C	-	+	+
	50	Leaf-napkin (leaves used to clean body)	-	+	C	+	-	C	C
	51	Leaf-dab (leaf dabbed on wound, examined)	-	+	+	-	-	C	-
	52	Leaf-groom (intense 'grooming' of leaves)	-	-	C	C	C	C	+
	53	Leaf-clip, mouth (rip parts off leaf, with mouth)	C	C	-	C	C	H	C
	54	Leaf-clip, fingers (rip leaf with fingers)	-	H	-	+	-	H	C
	55	Leaf-strip (rip leaves off stem, as threat)	+	-	H	+	-	H	-
	56	Leaf-squash (squash ectoparasite on leaf)	-	-	H	?	?	-	-
	57	Leaf-inspect (inspect ectoparasite on hand)	-	-	+	?	?	-	C
	58	Index-hit (squash ectoparasite on arm)	-	C	+	-	-	-	-
	59	Hand-clasp (clasp arms overhead, groom)	-	H	-	C	C	C	-
	60	Knuckle-knock (knock to attract attention)	+	C	H	C	C	-	-
	61	Branch din (bend, release saplings to warn)	-	-	-	-	-	-	-
	62	Branch-slap (slap branch, for attention)	C	C	-	+	-	-	C
	63	Stem pull-through (pull stems noisily)	C	-	+	H	-	H	-
	64	Shrub-bend (squash stems underfoot)	H	-	-	C	-	-	C
	65	Rain dance (slow display at start of rain)	-	H	C	C	C	C	H

A, Patterns absent at no site; B, patterns not achieving habitual frequencies at any site; C, patterns for which any absence can be explained by local ecological factors; D, patterns customary or habitual at some sites yet absent at others, with no ecological explanation. To facilitate comparison, behaviours are listed so that adjacent categories share broad functions; in Band D these are: 27–35, pounding actions; 36–40, fishing; 41–43, probing; 44 and 45, forcing; 46 and 47, comfort behaviour; 48 and 49, miscellaneous exploitation of vegetation properties; 50–57, exploitation of leaf properties; 58–59, grooming; 60–64, attention-getting. Sites (with subspecies, observation period in years by September 1998, site director): Bs, Bossou, Guinea (*verus*, 23, Y.S.); Ta, Tai Forest, Ivory Coast (*verus*, 23, C.B.); Go, Gombe, Tanzania (*schweinfurthii*, 38, J.G.); Ma, Mahale M-group, Tanzania (*schweinfurthii*, 30, T.N.); Mk, Mahale K-group (*schweinfurthii*, 18, T.N.); Kib, Kibale Forest, Uganda (*schweinfurthii*, 11, R.W.W.); Bd, Budongo Forest, Uganda (*schweinfurthii*, 8, V.R.). C, customary; H, habitual; +, present; -, absent; e, absent with ecological explanation; e?, ecological explanation suspected; (-), absent possibly because of inadequate observation; ?, answer uncertain (see text for full definitions). Branch din (behaviour 61) is allocated to band D because it is known to be customary at Lopé, Gabon (C.E.G.T.); behaviours 13, 15–17 and 22 are allocated to band B because they have been recorded at shorter-term sites (see Supplementary Information). For full definitions of all behaviours, see Supplementary Information.

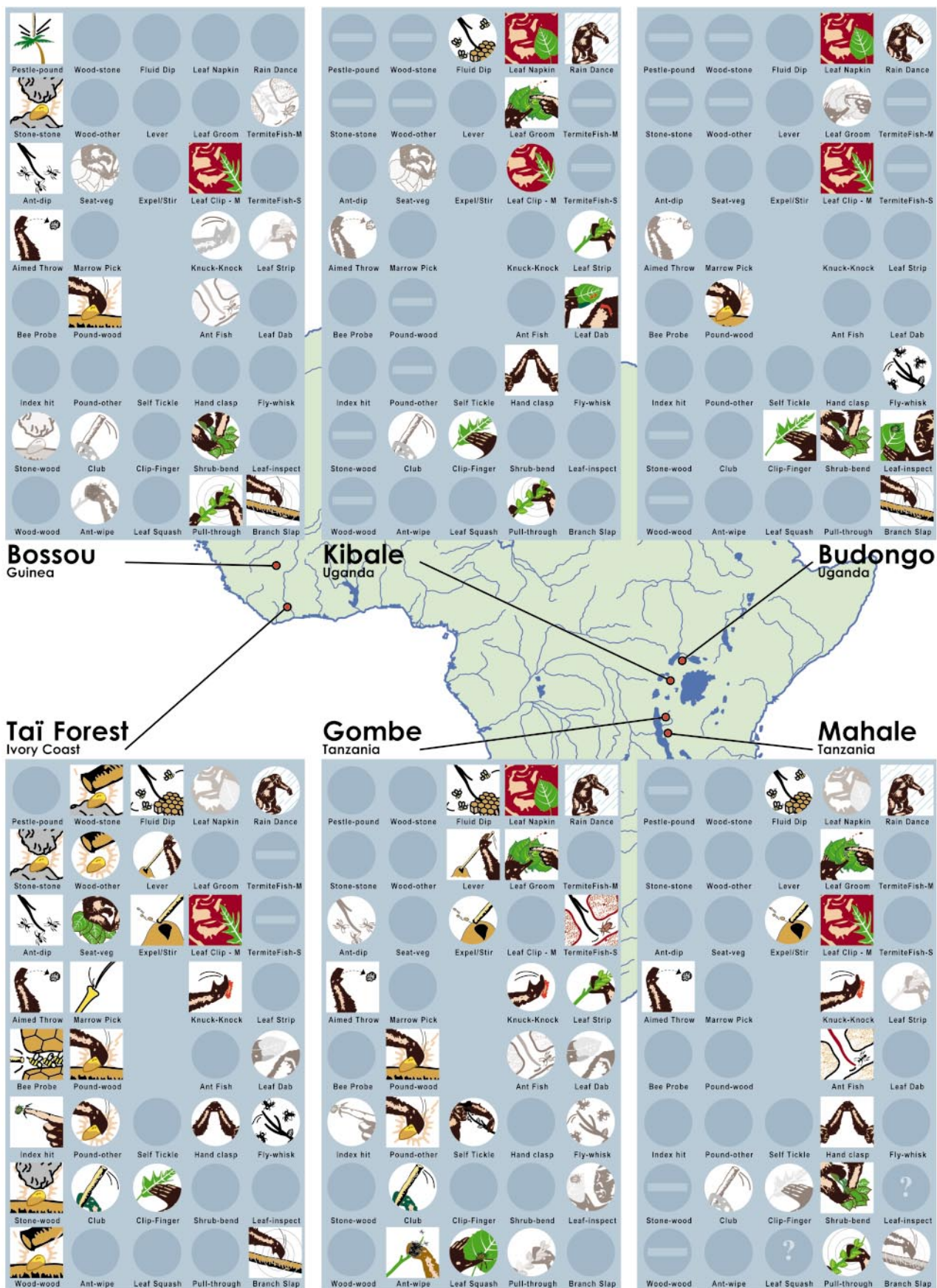


Figure 1 Distribution of behaviour patterns from band D in Table 1 across six African study sites. Behaviours are arranged in the 5x8 arrays to cluster those behaviours customary or habitual at each site, with clusters for westerly sites on the left of the array and clusters for easterly sites on the right. The secondary

Mahale site (K) is omitted. Colour icons, customary; circular icons, habitual; monochrome icons, present; clear, absent; horizontal bar, absent with ecological explanation; question mark, answer uncertain.

phase 1 were shown instead to be either customary or habitual in all communities (band A in Table 1). Second, 16 patterns failed to achieve habitual status in any community (band B in Table 1). The third class includes profiles in which all cases of absence are explicable by local conditions (band C in Table 1); just three cases were identified. Absence of algae-fishing can be explained by the rarity of algae, and any absence of ground night-nesting by high predator risk. Use of an additional stone to balance an anvil (anvil-prop) occurs only at Bossou, but it is not expected elsewhere because stone anvils are either not used or (at Tai) are embedded in the ground.

The remaining behaviours are absent at some sites but are customary or habitual at others (band D in Table 1). We have found 39 such behavioural variants, significantly more than previously suspected for chimpanzees^{1–6}. We know of no comparable variation in other non-human species, although no systematic study of this kind appears to have been attempted.

We arrive at a similar comparative conclusion when we examine the overall profiles of cultural variants in the different communities (Fig. 1). Some customary and habitual patterns are unique to certain communities, but others are shared between two or more communities (Table 2), so the clusters of variants that characterize each community are not mutually exclusive. Nevertheless, the profiles of each community (Fig. 1) are distinctively different, each with a pattern comprising many behavioural variants. These patterns vary as much between sites associated with the same subspecies (*verus* at Bossou and Tai in the west, and *schweinfurthii* at the four eastern sites) as between subspecies themselves. The only major difference between the western and eastern populations is that nut-cracking occurs only in the west, although the fact that this behaviour terminates abruptly at the Sassandra-N'Zo river within the range of the *verus* subspecies²¹ shows that it is culturally, rather than genetically, transmitted. The patterns in Fig. 1 can thus be seen to resemble those in human societies, in which differences between cultures are constituted by a multiplicity of variations in technology and social customs¹⁴. It remains to be shown whether chimpanzees are unique in this respect, or whether any other animal species, if studied in the same way, would reveal qualitatively similar patterns.

Other comparisons between human and non-human animal cultures have focused on the cognitive processes involved, arguing that if processes of human cultural transmission, such as imitative learning and teaching, are not found in animals, then culture in animals is merely an analogue of that in humans, rather than homologous with it^{22,23}. Our data agree with experimental studies that have shown that chimpanzees copy the methods used by others to manipulate and open artificial 'fruits' designed as analogues of wild foods^{24,25}. These experimental designs show differential copying of each of two quite different methods used to process the foods. Similarly, some of the differences between communities described here represent not only the contrast between habitual versus absent, but also the contrast between different versions of an otherwise similar pattern. Examples include cases of tool use, such as the two different methods of ant-dip (Table 1, items 39 and 40); in the first of these, a long wand is held in one hand and a ball of ants is wiped off with the other, whereas in the second method a short stick is held in one hand and used to collect a smaller number of ants, which are transferred directly to the mouth. Other examples occur in social behaviour, such as the variants used to deal with ectoparasites discovered during grooming, with leaf-squash, leaf-inspect and index-hit occurring in different communities (Table 1, items 56–58). It is difficult to see how such behaviour patterns could be perpetuated by social learning processes simpler than imitation, the most commonly suggested alternative to which is stimulus enhancement²⁶, in which the attention of an observer is merely drawn to a relevant item such as a stick. But this does not mean that imitation is the only mechanism at work. Experimental studies on the acquisition of tool-use and food-processing skills by both children and captive chimpanzees indicate that there is a complex

Table 2 Number of unique versus shared patterns that are either customary or habitual

	Bs	Ta	Go	Ma	Mk	Kib	Bd
Unique	1	8	3	0	1	1	1
Shared	8	16	13	11	9	9	8

Sites are abbreviated as in Table 1. Unique refers to customary or habitual patterns unique to the sites; shared refers to customary or habitual patterns shared with other sites. Frequencies exclude 'universal' behaviour patterns identified in band A of Table 1.

mix of imitation, other forms of social learning, and individual learning^{24,25,27–30}.

Our results show that chimpanzees, our closest sister-species, have rich behavioural complexity. However, although this study represents the definitive state of knowledge at present, we must expect that more extended study will elaborate on this picture. Every long-term study of wild chimpanzees has identified new behavioural variants. □

Received 24 March; accepted 11 May 1999.

- McGrew, W. C. & Tutin, C. E. G. Evidence for a social custom in wild chimpanzees? *Man* **13**, 234–251 (1978).
- Goodall, J. *The Chimpanzees of Gombe: Patterns of Behavior* (Harvard Univ. Press, Cambridge, Massachusetts, 1986).
- Nishida, T. *The Chimpanzees of the Mahale Mountains: Sexual and Life History Strategies* (Tokyo Univ. Press, Tokyo, 1990).
- McGrew, W. C. *Chimpanzee Material Culture: Implications for Human Evolution* (Cambridge Univ. Press, Cambridge, 1992).
- Sugiyama, Y. in *The Use of Tools by Human and Non-human Primates* (eds Berthelot, A. & Chavillon, J.) 175–187 (Clarendon, Oxford, 1993).
- Wrangham, R. W., McGrew, W. C., de Waal, F. B. M. & Heltne, P. G. (eds) *Chimpanzee Cultures* (Harvard Univ. Press, Cambridge, Massachusetts, 1994).
- Boesch, C. The emergence of cultures among wild chimpanzees. *Proc. Br. Acad.* **88**, 251–268 (1996).
- Bonner, J. T. *The Evolution of Culture in Animals* (Princeton Univ. Press, New Jersey, 1980).
- Mundinger, P. C. Animal cultures and a general theory of cultural evolution. *Ethol. Sociobiol.* **1**, 183–223 (1980).
- Lefebvre, L. & Palameta, B. in *Social Learning: Psychological and Biological Perspectives* (eds Zentall, T. & Galef, B. G. Jr) 141–164 (Erlbaum, Hillsdale, New Jersey, 1988).
- McGrew, W. C. Culture in non-human primates? *Annu. Rev. Anthropol.* **27**, 301–328 (1998).
- Marler, P. & Tamura, M. Song 'dialects' in three populations of white-crowned sparrows. *Science* **146**, 1483–1486 (1964).
- Catchpole, C. K. & Slater, P. J. B. *Bird Song: Themes and Variations* (Cambridge Univ. Press, Cambridge, 1995).
- Murdock, G. P. *Ethnographic Atlas* (Univ. Pittsburgh Press, Pittsburgh, 1967).
- Kroeber, A. L. & Kluckhohn, C. *Culture: A Critical Review of Concepts and Definitions* (Random House, New York, 1963).
- Bloch, M. Language, anthropology and cognitive science. *Man* **26**, 183–198 (1991).
- Nishida, T. in *Primate Societies* (eds Smuts, B. B., Cheney, D. L., Seyfarth, R. M., Wrangham, R. W. & Struhsaker, T. T.) 462–474 (Univ. Chicago Press, Chicago, 1987).
- Whiten, A. & Ham, R. On the nature of imitation in the animal kingdom: reappraisal of a century of research. *Adv. Study Behav.* **21**, 239–283 (1992).
- Imanishi, K. Identification: A process of enculturation in the subhuman society of *Macaca fuscata*. *Primates* **1**, 1–29 (1957).
- Huffman, M. in *Social Learning in Animals: The Roots of Culture* (eds Heyes, C. M. & Galef, B. G.) 267–289 (Academic Press, London, 1996).
- Boesch, C., Marchesi, P., Marchesi, N., Fruth, B. & Joulain, F. Is nut cracking in wild chimpanzees a cultural behaviour? *J. Hum. Evol.* **26**, 325–338 (1994).
- Galef, B. G. Jr The question of animal culture. *Hum. Nature* **3**, 157–178 (1992).
- Tomasello, M., Kruger, A. C. & Ratner, H. H. Cultural learning. *Behav. Brain Sci.* **16**, 495–552 (1993).
- Whiten, A., Cusance, D. M., Gomez, J.-C., Teixidor, P. & Bard, K. A. Imitative learning of artificial fruit-processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* **110**, 3–14 (1996).
- Whiten, A. Imitation of the sequential structure of actions by chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* **112**, 270–281 (1998).
- Spence, K. W. Experimental studies of learning and the mental processes in infra-human primates. *Psychol. Bull.* **34**, 806–850 (1937).
- Sumita, K., Kitahara-Frisch, J. & Norikoshi, K. The acquisition of stone tool use in captive chimpanzees. *Primates* **26**, 168–181 (1985).
- Tomasello, M., Davis, Dasilva, M., Camak, L. & Bard, K. Observational learning of tool-use by young chimpanzees. *Hum. Evol.* **2**, 175–183 (1987).
- Paquet, D. Discovering and learning tool-use for fishing honey by captive chimpanzees. *Hum. Evol.* **7**, 17–30 (1992).
- Nagell, K., Olguin, K. & Tomasello, M. Processes of social learning in the tool use of chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *J. Comp. Psychol.* **107**, 174–186 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature. An extended graphical database (unrefereed) of this material is also available (<http://chimp.st-and.ac.uk/cultures>).

Acknowledgements. We thank T. Matsuzawa, G. Yamakoshi, H. Boesch, D. A. Collins, S. Kamenya, H. Matama, H. Mkonzo, E. Mpongo, J. Salala, M. Huffman, M. Kasagula, R. Nyundo, S. Uehara, K. Arnold, C. Assersohn, K. Fawcett, J. Kakura, Z. Kiwele, G. Muhumuzi, N. Newton-Fisher, P. Pebsworth, E. Stokes, J. Tinka, A. Arcadi, C. Katongole, G. Isabirye-Basuta, F. Mugurusi, M. Muller and M. Wilson for contributions to the database; D. A. Collins, D. I. Perrett and P. J. B. Slater for advice on the manuscript; and S. Smart for the graphics of Fig. 1.

Correspondence and requests for materials should be addressed to A.W.

Jasmonate-inducible plant defences cause increased parasitism of herbivores

Jennifer S. Thaler

Department of Entomology, University of California, Davis, One Shields Avenue, Davis, California 95616-8584, USA

In many plants, defence systems against herbivores are induced through the octadecanoid pathway^{1,2}, which may also be involved in recruiting natural enemies of herbivores³. This pathway can be induced by treating plants with jasmonic acid⁴ or by natural herbivory, and increases resistance against herbivorous insects in tomato plants⁵, in part by causing production of toxic and antinutritive proteinase inhibitors and oxidative enzymes^{6–8}. Herbivore-infested tomato plants release increased amounts of volatiles⁹ and attract natural enemies of the herbivores¹⁰, as do other plants^{11–15}. The octadecanoid pathway may regulate production of these volatiles, which attract host-seeking parasitic wasps^{16,17}. However, plant resistance compounds can adversely affect parasitoids as well as herbivores¹⁸. It is unclear whether the combination of increased retention and/or attractiveness of parasitic wasps to induced plants and the adverse effects of plant defence compounds on both caterpillars and parasitoids results in a net increase in parasitization of herbivores feeding on induced plants. Here I show that inducing plants with jasmonic acid increases parasitism of caterpillar pests in an agricultural field twofold. Thus, elicitors of plant resistance may become useful in agriculture.

In the tomato growing region of the Central Valley of California, I stimulated the octadecanoid pathway by spraying eight-week-old tomato plants (*Lycopersicon esculentum* var. Ace) with 0.5 mM (circa 1.5 μ mol per plant) jasmonic acid. One-hundred-and-three plots containing 4–6 plants each were randomly divided into plants treated with jasmonic acid (induced plants) and plants sprayed with water and acetone carrier (control). The number of naturally occurring herbivores on induced plants and control plants was not manipulated.

I confirmed that the octadecanoid pathway was induced by measuring polyphenol oxidase activity (a putative defence) in the induced and control plants four days after the plants were sprayed (mean change in optical density per gram per minute $\pm$ s.e.; control: 6.739 ± 0.884 ; induced: 19.839 ± 2.389 ; analysis of variance: $F_{1,38} = 26.454$, $P < 0.001$) (methods in ref. 5). Next I examined the effects of induction on the plant–herbivore–natural enemy interaction in a field with natural infestations of *Hyposoter exiguae* and *Spodoptera exigua*. *H. exiguae* is an endoparasitic wasp and an important killer of the agronomic lepidopteran pest, *S. exigua*¹⁹. Three weeks after the field was sprayed, the number of naturally occurring *H. exiguae* pupae, which had developed in native, naturally occurring *S. exigua* caterpillars, were counted on one plant in each plot. There were twice as many *H. exiguae* pupae on induced plants as on controls (Kruskal–Wallis: $n = 103$, $\chi^2 = 5.12$, d.f. = 1, $P = 0.024$; Fig. 1a).

To confirm that plant traits per se were responsible for this increased parasitism, I placed sentinel caterpillars underneath the canopy of jasmonate-induced and control plants in the field. The sentinel *S. exigua* caterpillars were reared on an artificial diet and then placed in cups in the field for 24 h; these sentinel caterpillars could not feed on the plant, but were still accessible to foraging parasitoids (see Methods). There was 37% more parasitism in caterpillars associated with induced plants than in those associated with controls (Kruskal–Wallis: $n = 103$, $\chi^2 = 4.57$, d.f. = 1,

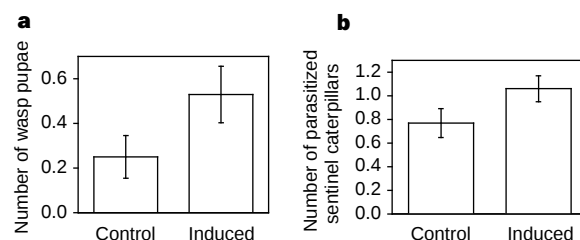


Figure 1 The effects of jasmonate-induced plant responses on the parasitism of caterpillars. **a**, The number of *Hyposoter exiguae* wasp pupae found on the foliage of induced and control plants. The number of wasp pupae is a measure of the number of parasitized caterpillars. **b**, The mean number of parasitized sentinel *Spodoptera exigua* caterpillars (reared on an artificial diet and never exposed to plants) from underneath the canopies of induced and control plants. Bars indicate mean $\pm$ s.e.

$P = 0.032$; Fig. 1b). Based on these findings, expression of the octadecanoid pathway appears to result in increased attractiveness and/or retention of parasitoids to plants and increased mortality for injurious herbivorous insects, irrespective of herbivore quality, developmental stage or quantity.

Induced plant resistance can also negatively affect parasitic wasps. In a series of controlled laboratory experiments, wasps were reared on caterpillars to assess the indirect effects of plant resistance on parasitoid success through altered host quality. In the first experiment, caterpillars from induced and control foliage were matched for age at the time of parasitism by naive, mated female wasps (see Methods). In the second experiment, caterpillars were matched for body mass. In both experiments, parasitoid performance was reduced on caterpillar hosts grown on induced compared to those grown on control foliage (Fig. 2a, b; multivariate analysis of variance; same age, different mass: Wilk's $\lambda = 0.204$, $F_{3,12} = 15.613$, $P < 0.001$; same mass, different age: Wilk's $\lambda = 0.537$, $F_{3,26} = 7.487$, $P = 0.001$). Univariate analyses showed that egg-to-pupa development time was increased and pupal mass was decreased on induced plants (same age, different mass: egg-to-pupa: $F_{1,14} = 9.022$, $P = 0.009$; pupal mass: $F_{1,14} = 31.798$, $P < 0.001$; same mass, different age: egg-to-pupa: $F_{1,28} = 20.931$, $P < 0.001$; pupal mass: $F_{1,28} = 5.268$, $P = 0.029$); whereas pupa-to-adult development time was not different for wasps reared in caterpillars on induced and control foliage (same age, different mass: $F_{1,14} = 1.491$, $P = 0.242$; same mass, different age: $F_{1,28} = 0.361$, $P = 0.553$). There was no effect of treatment on wasp survival (same age, different mass: $n = 23$, $\chi^2 = 0.140$, d.f. = 1, $P = 0.708$; same mass, different age: $n = 41$, $\chi^2 = 0.211$, d.f. = 1, $P = 0.646$). When tested in the absence of plants, the parasitoids could not differentiate between equal-sized caterpillars grown on induced and control foliage (χ^2 : $n = 28$, $\chi^2 = 0.142$, $P > 0.75$). Wasp attraction to jasmonic acid itself was tested by simultaneously displaying pieces of paper sprayed with jasmonic acid or acetone control solutions in laboratory cages and observing which paper female wasps oriented to first in five minutes. Wasps were not directly attracted to or repelled by jasmonic acid compared to acetone controls (χ^2 : 60 trials, 14 = no choice, 20 = jasmonic acid, 26 = control, $n = 46$, $\chi^2 = 0.783$, $P = 0.345$).

Plant resistance to herbivores mediated by the octadecanoid pathway has both positive and negative influences on parasitoids. The net effect of this interaction on herbivore mortality was a doubling of the number of naturally occurring parasitized caterpillars on induced plants compared to controls. Production of volatiles that influence attraction of natural enemies may be vital to the plant's defence system, overcoming the negative influences of induction on wasp performance. There are specific interactions between herbivores and plants that can result in a varied blend of volatiles being released from plants, which selectively attract parasitoids^{15,20}. The parasitoids in this study could have utilized

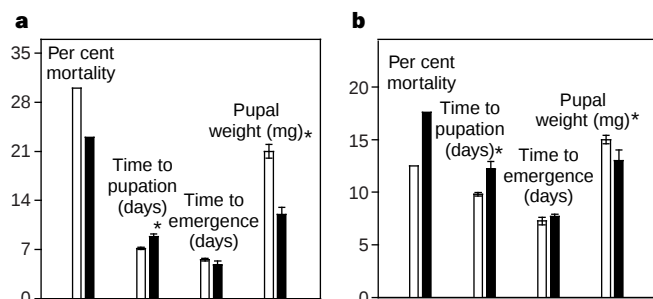


Figure 2 Indirect effects of jasmonate-induced plant responses on parasitoid performance due to changes in the caterpillar host. **a**, The per cent mortality, time to pupation, time to adult emergence and pupal mass of *H. exiguae* wasps reared in *S. exigua* caterpillars feeding on induced (black bars) and control (white bars) foliage which were the same age but different mass. **b**, The per cent mortality, time to pupation, time to adult emergence and pupal mass of wasps reared in caterpillars feeding on induced and control foliage which were the same mass but different ages. Asterisks indicate measures of parasitoid performance which were statistically significant at $P < 0.05$. Bars indicate mean $\pm$ s.e.

either herbivore-specific or nonspecific plant cues. Because jasmonic acid was used to induce plants and attract parasitoids of caterpillars, activation of the octadecanoid pathway in tomato plants either specifically stimulates the same plant responses as *S. exigua* feeding or stimulates non-herbivore-specific cues.

There are alternative hypotheses to the increased attraction or retention of wasps to induced plants that could explain the increased rate of parasitism found in this study. One hypothesis is that herbivores feeding on induced plants have a weakened immune system, increasing their vulnerability to parasitoid attack. My data do not support this hypothesis (Fig. 2), as the same proportion of wasps successfully completed development in caterpillars from control and induced foliage. Alternatively, caterpillars on induced plants could have reduced growth rates, leading to an increase in the period of time when the caterpillars can be successfully parasitized (slow-growth-high-mortality hypothesis^{21,22}). *S. exigua* larvae feeding on induced plants have reduced growth rates compared to larvae feeding on control plants⁵, and are primarily parasitized in the late first and second instars. In additional experiments, caterpillars were exposed to parasitoids in replicated field cages with either control or induced plants. This experiment examined the effect of induced resistance when parasitoid attraction was reduced because the parasitoids were confined within the cage (see Methods). The number of caterpillars in the cages with control and induced plants was equalized, but the caterpillars were of different sizes because they grew more slowly on the induced plants. Herbivore mortality due to parasitism was higher on control than induced plants (induction: $F_{1,10} = 0.002$, $P = 0.966$ (no significant effect because the number of herbivores in induced and control cages was manually equalized); parasitism: $F_{1,10} = 5.514$, $P = 0.041$; induction $\times$ parasitism: $F_{1,10} = 5.209$, $P = 0.046$). This finding contradicts the key prediction of the slow-growth-high-mortality hypothesis, that caterpillars should experience higher parasitism on induced plants compared to controls due to prolonged development and longer exposure to parasitoids. Thus, the increased parasitism in this study (Fig. 1a) is not likely to be due to prolonged development of the caterpillars. That sentinel caterpillars were parasitized more often near induced plants than control plants (Fig. 1b) reinforces the assertion that factors independent of herbivore quality are important for increased parasitism on induced plants.

Herbivory and application of jasmonates to plants are associated with increased production of volatile compounds¹². In tomato plants, spider-mite damage and jasmonate spray result in similar, but not identical induction of volatile compounds. Both result in a large increase in the proportion of methyl salicylate and terpenoids found in the volatile blend emitted; minor inducible constituents

include (Z)-3-hexen-1-ol, a C-6 leafy green volatile^{9,16}. The localized production of C-6 leafy green volatiles may be due to the activation of the octadecanoid cascade through the action of lipoxygenase, an enzyme involved in the induction of both foliar and volatile phytochemicals^{23–25}. In addition, it has been proposed that volicitin, a compound found in insect oral secretions of plant and herbivore origin, may interact with the octadecanoid pathway and result in increased production of volatile compounds in corn plants that are attractive to natural enemies²⁶. The multiplicity of effects provided by the octadecanoid pathway indicates that plants coordinate an interrelated, complex system of defence.

The octadecanoid pathway contributes to plant resistance both by directly killing herbivores and by enhancing the action of natural enemies of herbivores. In tomato plants, jasmonate-induced resistance resulted in reduced abundance of herbivores in three feeding guilds, including leaf chewers, phloem feeders and cell content feeders (J.S.T., M. J. Stout, R. Karban and S. S. Duffey, unpublished results). The finding that jasmonate-induced resistance results in increased mortality of herbivores due to parasitism is potentially of great importance in pest control. Future pest management strategies will rely on the use of combinations of strategies; these data point towards conditions where induced plant resistance and biological control may provide compatible and potentially synergistic control of herbivores²⁷. □

Methods

Sentinel caterpillars were reared on an artificial diet (Southland Products) for five days and then placed in the field in clear plastic cups without lids. The cups did not contact the plant's stems or leaves but were within the canopy. Ten caterpillars were placed in each cup. Because the cups are slippery, the caterpillars could not crawl out of the cups, but the parasitoids were able to enter the cups. The experiment was started 10 days after the field was sprayed with jasmonic acid. After being exposed to parasitism for 24 h the caterpillars were brought back to the lab and reared for parasites.

Rearing caterpillars on induced or control foliage for their entire larval period tested the indirect effects of plant resistance on parasitoid success. In the same mass, different age experiment, the caterpillars were reared for 7 days on control foliage and 10 days on induced foliage, parasitized, and returned to the appropriate foliage. In the same age, different mass experiment, all caterpillars were parasitized on day 8.

The design of the field cage experiment was a 2×2 factorial, with control and induced plants in cages (cage size: 80 m²) crossed with and without parasitoids augmented. Colony-reared *S. exigua* larvae were added to each cage, and allowed to feed for one week, after which the number of caterpillars in control or induced cages was equalized. At this point, two mated female parasitoids were added twice a week to the cages receiving parasitism until all of the caterpillars were beyond the stage where parasitism can occur.

Received 19 March; accepted 28 April 1999.

- Karban, R. & Baldwin, I. T. *Induced Responses to Herbivory* (Univ. of Chicago Press, Chicago, Illinois, 1997).
- Meyer, A., Miersch, O., Buttner, C., Dathe, W. & Sembdner, G. Occurrence of the plant growth regulator jasmonic acid in plants. *J. Plant Growth Regul.* 3, 1–8 (1984).
- Farmer, E. E. New fatty acid based signals: a lesson from the plant world. *Science* 276, 912–913 (1997).
- McConn, M., Creelman, R. A., Bell, E., Mullet, J. E. & Browse, J. Jasmonate is essential for insect defense in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* 93, 5473–5477 (1996).
- Thaler, J. S., Stout, M. J., Karban, R. & Duffey, S. S. Exogenous jasmonates simulate insect wounding in tomato plants, *Lycopersicon esculentum*, in the laboratory and field. *J. Chem. Ecol.* 22, 1767–1781 (1996).
- Broadway, R. M., Duffey, S. S., Pearce, G. & Ryan, C. A. Plant proteinase inhibitors: a defense against herbivorous insects? *Entomol. Exp. Appl.* 41, 33–38 (1986).
- Felton, G. W., Donato, K., del Vecchio, R. J. & Duffey, S. S. Activation of plant foliar oxidases by feeding reduces nutritive quality of foliage for noctuid herbivores. *J. Chem. Ecol.* 15, 2667–2694 (1989).
- Orozco-Cardenas, M., McGurl, B. & Ryan, C. A. Expression of an antisense prosystemin gene in tomato plants reduces resistance toward *Manduca sexta* larvae. *Proc. Natl Acad. Sci. USA* 90, 8273–8276 (1993).
- Dicke, M., Takabayashi, J., Posthumus, M. A., Schutte, C. & Krips, O. E. Plant-phytoseid interactions mediated by herbivore-induced plant volatiles: variation in production of cues and in responses of predatory mites. *Exp. Appl. Acarol.* 22, 311–333 (1998).
- Dicke, M. & Sabelis, M. W. How plants obtain predatory mites as bodyguards. *Netherlands J. Zool.* 38, 148–165 (1988).
- Turlings, T. C. J., McCall, P. J., Alborn, H. T. & Tumlinson, J. H. An elicitor in caterpillar oral secretions that induces corn seedlings to emit chemical signals attractive to parasitic wasps. *J. Chem. Ecol.* 19, 411–425 (1993).

12. Hopke, J., Donath, J., Blechert, S. & Boland, W. Herbivore-induced volatiles: the emission of acyclic homoterpenes from leaves of *Phaseolus lunatus* and *Zea mays* can be triggered by a β -glucosidase and jasmonic acid. *FEBS Lett.* **352**, 146–150 (1994).
13. Drukker, B., Scutareanu, P. & Sabelis, M. W. Do anthocorid predators respond to synomones from *Psylla*-infested pear trees under field conditions? *Entomol. Exp. Appl.* **77**, 193–203 (1995).
14. Shimoda, T., Takabayashi, J., Ashihara, W. & Takafuji, A. Response of predatory insect *Scotolophts takahashii* toward herbivore-induced plant volatiles under laboratory and field conditions. *J. Chem. Ecol.* **23**, 2033–2048 (1997).
15. De Moraes, C. M., Lewis, W. J., Paré, P. W., Alborn, H. T. & Tumlinson, J. H. Herbivore-infested plants selectively attract parasitoids. *Nature* **393**, 570–572 (1998).
16. Boland, W., Hopke, J., Donath, J., Nuske, J. & Bublitz, F. Jasmonic acid and coronatin induce odor production in plants. *Angew. Chem. Int. Ed. Engl.* **34**, 1600–1602 (1995).
17. Alborn, H. T. *et al.* An elicitor of plant volatiles from beet armyworm oral secretion. *Science* **276**, 945–949 (1997).
18. Campbell, B. C. & Duffey, S. S. Tomatine and parasitic wasps: Potential incompatibility of plant-antibiosis with biological control. *Science* **205**, 700–702 (1979).
19. Strand, L. L. *Integrated Pest Management for Tomatoes* (Division of Agriculture and Natural Resources Communication Services—Publications, Oakland, 1998).
20. Takabayashi, J. & Dicke, M. Plant-carnivore mutualism through herbivore-induced carnivore attractants. *Trends Plant Sci.* **1**, 109–113 (1996).
21. Clancy, K. M. & Price, P. W. Rapid herbivore growth enhances enemy attack: sublethal plant defenses remain a paradox. *Ecology* **68**, 733–737 (1987).
22. Benrey, B. & Denno, R. F. The slow-growth-high-mortality hypotheses: a test using the cabbage butterfly. *Ecology* **78**, 987–999 (1997).
23. Stout, M. J., Workman, K. V. & Duffey, S. S. Identity, spatial distribution, and variability of induced chemical responses in tomato plants. *Entomol. Exp. Appl.* **79**, 255–271 (1996).
24. Bell, E. & Mullet, J. E. Characterization of an *Arabidopsis* lipoxygenase gene responsive to methyl jasmonate and wounding. *Plant. Physiol.* **103**, 1133–1137 (1993).
25. Siedow, J. N. Plant lipoxygenase: structure and function. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **42**, 145–188 (1991).
26. Paré, P. W., Alborn, H. T. & Tumlinson, J. H. Concerted biosynthesis of an insect elicitor of plant volatiles. *Proc. Natl Acad. Sci. USA* **95**, 13971–13975 (1998).
27. Dicke, M., Sabelis, M. W., Takabayashi, J., Bruin, J. & Posthumus, M. A. Plant strategies of manipulating predator-prey interactions through allelochemicals: prospects for application to pest control. *J. Chem. Ecol.* **16**, 3091–3118 (1990).

Acknowledgements. I thank C. Wardlaw and C. Black for help in the field; M. Stout and C. De Moraes for discussion; and A. Agrawal, R. Karban, M. Dicke, J. Rosenheim, L. Adler, C. Black, R. Kimura and J. Granett for comments on the manuscript. This research was supported by the USDA-NRI.

Correspondence and requests for materials should be addressed to J.S.T. (e-mail: jsthaler@ucdavis.edu).

Auditory cortical responses in the cat to sounds that produce spatial illusions

Li Xu, Shigeto Furukawa & John C. Middlebrooks

Kresge Hearing Research Institute, University of Michigan, 1301 E. Ann St, Ann Arbor, Michigan 48109-0506, USA

Humans and cats can localize a sound source accurately if its spectrum is fairly broad and flat^{1–3}, as is typical of most natural sounds. However, if sounds are filtered to reduce the width of the spectrum, they result in illusions of sources that are very different from the actual locations, particularly in the up/down and front/back dimensions^{4–6}. Such illusions reveal that the auditory system relies on specific characteristics of sound spectra to obtain cues for localization⁷. In the auditory cortex of cats, temporal firing patterns of neurons can signal the locations of broad-band sounds^{8–9}. Here we show that such spike patterns systematically mislocalize sounds that have been passed through a narrow-band filter. Both correct and incorrect locations signalled by neurons can be predicted quantitatively by a model of spectral processing that also predicts correct and incorrect localization judgements by human listeners⁶. Similar cortical mechanisms, if present in humans, could underlie human auditory spatial perception.

Resonance in the cavities of the external ears, combined with diffraction by the head, act to enhance particular frequencies in an arriving sound and to cancel other frequencies. The filter properties of the head and ears vary according to the direction of incidence of a sound, so details of the frequency spectra at the eardrums can reveal the direction of the source. Interpretation of spectral cues for localization is confounded when sound sources are filtered experimentally^{2–6}. For instance, in a study in which stimuli were restricted to 1/6-octave frequency bands, human listeners reported a compel-

ling illusion of a sound source that deviated in vertical or front/back location from the actual source⁶. Regardless of the stimulus location, a typical listener in that study consistently reported hearing sound sources high and in front when the centre frequency of the stimulus spectrum was 6 kHz, low and generally to the front when the centre frequency was 8 kHz, and low and generally to the rear when the centre frequency was 10 kHz. An acoustical model based on measurements of the filter characteristics of each listener's ears successfully predicted the listener's locations judgments. Listeners appeared to interpret narrow-band stimulus spectra as if they were broad-band spectra that had been filtered by their own external ears. In cats, two recent studies^{2–3} have demonstrated localization behaviour that was generally consistent with human results. Like humans, cats localized broad-band sounds accurately^{2–3} and made large localization errors when presented with 1/6-octave narrow-band sounds³. Unlike humans, the cats' responses did not vary systematically with stimulus centre frequency. That negative result, however, might reflect the response measure that was used in the cat experiments, which constrained the cats' responses to a limited frontal area. Our localization model predicts that the narrow-band stimuli that were tested in that study would have produced illusory targets high above or behind the cats.

In this study, we recorded unit activity from cortical area A2 in cats that were anaesthetized with α -chloralose⁸. Area A2 was chosen because neurons in that area respond to broad ranges of frequency and because elevation sensitivity tends to be greater there than in the adjacent anterior ectosylvian sulcus area⁹ or in area A1 (unpublished observations). We used silicon-substrate thin-film probes that permitted simultaneous recording of unit activity at up to 16 cortical sites¹⁰. Custom spike-sorting software was used to isolate single units or small clusters of units. Each sound was presented in an anechoic room from one of 14 loudspeakers. The loudspeakers were located in the vertical midline, all at a distance of 1.2 m from the centre of the cat's head. We use the term elevation to refer to a continuous angular dimension measured relative to the frontal horizontal plane and extending up and over the head to the rear. Stimulus locations ranged in elevation in 20° increments from –60° (below the horizontal plane in front), through +90° (straight overhead) to +200° (20° below the horizontal plane in the rear). Stimuli consisted of filtered gaussian noise bursts, 80 ms long. Sound pressures were varied in 5- or 10-dB steps between 20 and 40 dB above neural thresholds. Broad-band filter spectra were flat and extended from 1 to 30 kHz. Narrow-band filters had 1/6-octave-wide flat centres, skirts that fell off at 120 dB per octave and centre frequencies that ranged from 4 to 18 kHz. We recorded from 389 single units and small clusters of units in area A2 of eight cats. When tested with tonal stimuli, most of the units responded to frequencies in the upper half of the cat's audible range, and all showed frequency bandwidths of an octave or more at levels 40 dB above threshold.

Neurons typically responded to broad-band noise bursts with one or a few spikes locked to the onset of the stimulus. Spike probabilities and spike latencies varied according to the elevation of the sound source. As in previous studies^{8,9}, we computed estimates of spike density functions (spike patterns) by forming averages of multiple samples of eight responses to each stimulus; samples were drawn with replacement. Spike patterns drawn from odd-numbered trials were used to train an artificial neural network; we then used the trained network to estimate sound-source elevations by classifying spike patterns drawn from even-numbered trials. Figure 1a shows the network estimates of source elevations based on the spike patterns of one cortical neuron. In this example of broad-band stimulus conditions, responses generally clustered around the dashed line that represented perfect performance.

The accuracy of elevation estimates under broad-band conditions, as represented by the medians of unsigned errors, ranged from 19.9° to around the chance-performance level of 65°. Responses to narrow-band stimuli were analysed for the half of

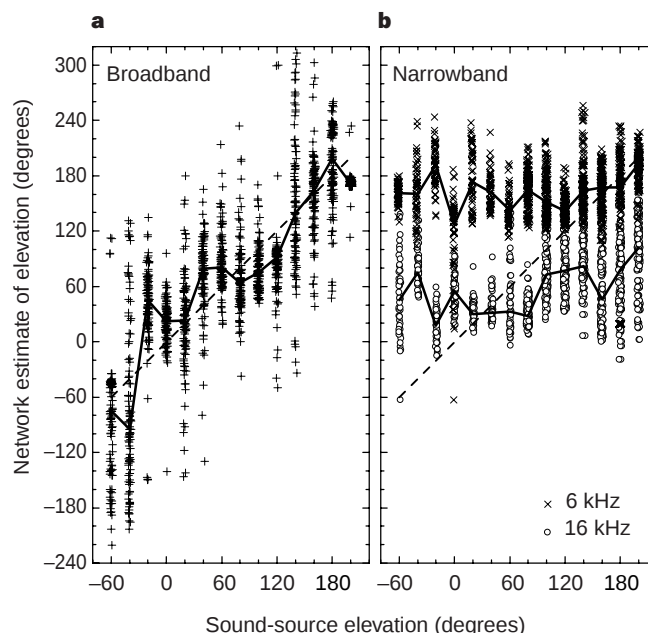


Figure 1 Artificial-neural-network analysis of spike patterns from a single unit (9806C16). The network was trained to estimate stimulus elevations based on spike patterns elicited by broad-band sounds that varied in 5-dB steps from 20 to 40 dB above neural threshold. The trained network was used to classify spike patterns elicited by broad-band (**a**) or narrow-band (**b**) sounds. Each symbol represents the network estimate of sound-source elevation based on an average of eight responses from one unit. The samples of eight responses were drawn with replacement. The dashed line indicates perfect correspondence between stimulus location and the network estimate. The solid line shows the median

value at each source location. **a**, Classification of responses to broad-band sounds. Stimulus levels were 20 dB above neural threshold. The median error across all source elevations was 31.5°. **b**, Classification of responses to 1/6-octave noise bands centred at 6 and 16 kHz. Stimulus levels were 40 dB above threshold. Unit responses to narrow-band sounds varied with stimulus centre frequency but were largely insensitive to sound-source elevation. When artificial neural networks were trained and tested with responses to narrow-band sound, performance was consistently worse than for broad-band conditions, generally around chance levels.

the unit sample (194 units, median errors $\leq 50.4^\circ$) that showed the most accurate elevation coding; complete data for narrow-band centre frequencies between 6 and 18 kHz were obtained for 112 to 146 units, depending on the frequency. For each unit, we trained an artificial neural network with spike patterns from broad-band conditions, then used the trained network to classify spike patterns that were elicited by 1/6-octave band-passed stimuli at various centre frequencies. We regard this procedure as a physiological analogue of requiring a psychophysical listener to report the apparent location of a narrow-band sound. Neuronal elevation sensitivity was markedly reduced when narrow-band stimuli were used. Figure 1b shows the network estimates of source elevations based on neuronal responses to noise bands centred at 6 and 16 kHz. Like the responses of human listeners in narrow-band localization tasks, unit responses were largely insensitive to source elevation and tended to vary according to the centre frequencies of the stimulus spectra. In this example, the network localized the 6-kHz stimulus as up and to the rear and the 16-kHz stimulus as up and in front. In this 16-kHz example, unit responses to front targets differed somewhat from those to rear targets, but elevation estimates were far from accurate.

To implement a model that related sound localization to external-ear acoustics⁶, we measured the filter characteristics of the external ears of the cats that were studied physiologically. The filter characteristics were represented by directional transfer functions (DTFs)^{9,11}. DTFs were measured by recording from the ear canals with microphones as a broad-band sound source was varied systematically in location. DTFs were similar among cats, in that upward elevation shifts of the sound source generally resulted in upward frequency shifts of spectral peaks and notches. Nevertheless, there was variability in the details of DTFs among cats that was at least qualitatively similar to the variability observed among human subjects¹². The acoustical model involved a computation of the

differences between stimulus spectra and DTFs (Fig. 2a). For each of the narrow-band stimuli and the DTF measured at each elevation, we subtracted the amplitude spectrum (gain) of the DTF from the stimulus spectrum, then computed the variance of the resulting difference distribution. That difference metric is referred to here as the spectral difference. In our previous psychophysical work⁶, listeners' localization judgements tended to fall at elevations for which their DTFs were most similar to the stimulus spectrum. Similarly, we predicted that the responses of cortical neurons would signal elevations that minimized the spectral difference between the stimulus spectrum and the cats' DTFs. Figure 2b and c shows the spectral differences computed for two cats for a range of stimulus centre frequencies. In these plots, each vertical column of coloured squares represents the spectral differences computed between a stimulus of a particular centre frequency and the DTFs at various elevations. The differences between the two panels are a result of differences between cats in the frequencies of components of their DTFs.

The plus signs in Fig. 2b and c represent artificial-neural-network estimates of elevation based on the responses of one neuron to narrow-band stimuli at various centre frequencies. The plus signs tend to fall on or near the yellow squares that indicate small spectral differences, which would be the predicted responses in a psychophysical task. The differences in cortical responses between Fig. 2b and 2c presumably reflect differences in acoustical input that resulted from differences in DTFs.

We tested the correspondence between cortical responses and model predictions by ranking spectral differences at each stimulus centre frequency (each vertical column in plots like Fig. 2b and c) for each cat. Low ranks indicate small spectral differences. On each physiological trial, we scored the rank of the elevation at which the network estimate fell. The median rank was recorded for each centre frequency for each of the units. Figure 3 shows the distribution of

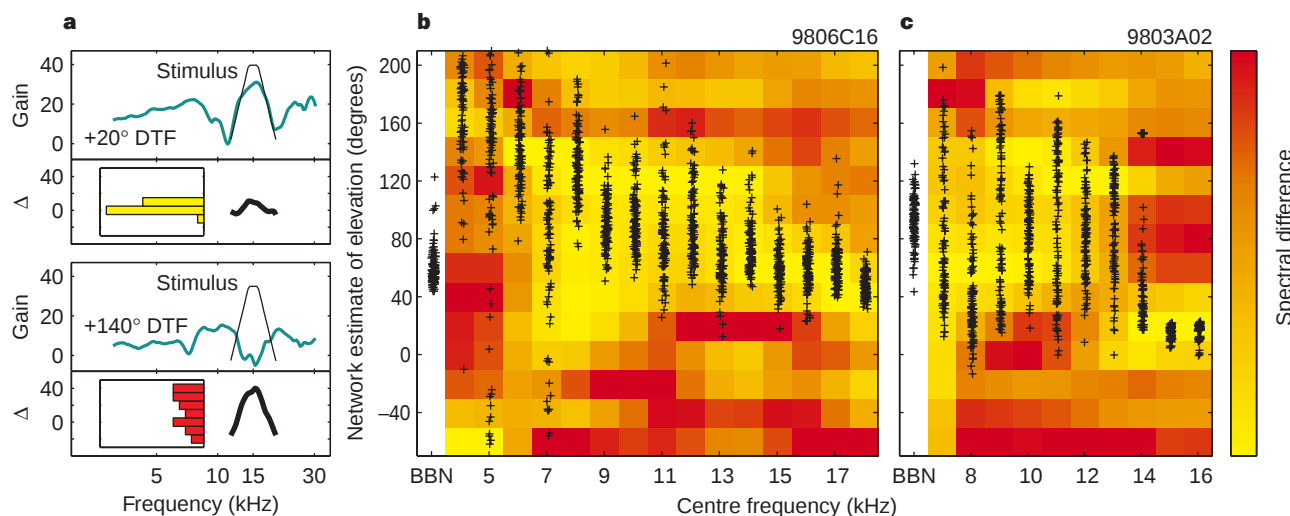


Figure 2 Spectral differences and network classification of cortical responses to narrow-band sounds. **a**, Spectral differences represent the differences between stimulus filter functions (thin black lines) and DTFs (green lines). The examples show a narrow-band stimulus centred at 15 kHz and DTFs for elevations of +20° and +140°. Filter functions and DTFs were expressed as gain in dB. Spectra were subtracted, frequency component by frequency component, to obtain difference spectra (Δ , thick black lines). The spectral difference metric was given by the variance of the resulting difference distribution (yellow and red histograms). **b**, Spectral differences and network classifications for unit 9806C16. Spectral

differences are represented by colours. Values at each stimulus centre frequency were scaled to span the full colour range. Light yellow indicates a small spectral difference, which indicates a prediction of a high likelihood of a response at the elevation. Symbols indicate network classifications of cortical responses to narrow-band stimuli. All sound sources were located at +80° elevation. The vertical column labelled BBN indicates the network classification of cortical responses to a broad-band sound presented from +80° elevation. **c**, Spectral variances and network classifications for unit 9803A02. Other conventions as in **b**.

median ranks across all units. Low-numbered ranks indicate good model performance, and chance performance would produce ranks around 7.5. Model performance was quite good for most units for centre frequencies between 11 and 16 kHz. Performance approached chance levels at lower and higher frequencies. It has been shown that cats can localize in elevation fairly accurately when tested with noise-band stimuli that are restricted to 10–14 kHz, whereas performance is near chance when stimuli are restricted to lower or higher frequencies².

Our results show successful prediction of cat neuronal signalling by a model that was derived to account for human psychophysical performance. Neurons signal accurately the elevations of broad-band sounds that humans localize accurately. In addition, in

response to narrow-band sounds, neurons signal incorrect locations that can be predicted by a model that predicts incorrect human localization. We would like to interpret this correspondence between psychophysical and physiological results as evidence that cortical responses analogous to those recorded in the cat could underlie human perception. That interpretation, of course, involves a conceptual jump that crosses the gap between species, cat to human, and crosses the gap across levels, from neurons to perception. Although procedural differences prevent an exact comparison between psychophysical results in humans and cats, particularly in regard to narrow-band stimuli, sound localization appears at least qualitatively similar between species. The jump across levels, from neurons to perception, is a general problem in neuroscience that cannot be solved by this study. Nevertheless, the results of this study do offer strong support for the conclusion that neuronal elevation sensitivity in cats and spatial hearing in humans obey similar computational principles. Neurons like those that we record in that cat auditory cortex, if present in humans, could play a key role in auditory spatial perception. □

Received 16 March; accepted 21 April 1999.

1. Makous, J. C. & Middlebrooks, J. C. Two-dimensional sound localization by human listeners. *J. Acoust. Soc. Am.* **87**, 2188–2200 (1990).
2. Huang, A. Y. & May, B. J. Sound orientation behavior in cats. II. Mid-frequency spectral cues for sound localization. *J. Acoust. Soc. Am.* **100**, 1070–1080 (1996).
3. Populin, L. C. & Yin, T. C. Behavioral studies of sound localization in cat. *J. Neurosci.* **18**, 2147–2160 (1998).
4. Blauert, J. Sound localization in the median plane. *Acustica* **22**, 205–213 (1969–1970).
5. Hebrank, J. & Wright, D. Spectral cues used in the localization of sound sources on the median plane. *J. Acoust. Soc. Am.* **56**, 1829–1834 (1974).
6. Middlebrooks, J. C. Narrow-band sound localization related to external ear acoustics. *J. Acoust. Soc. Am.* **92**, 2607–2624 (1992).
7. Middlebrooks, J. C. & Green, D. M. Sound localization by human listeners. *Annu. Rev. Psychol.* **42**, 135–159 (1991).
8. Middlebrooks, J. C., Xu, L., Eddins, A. C. & Green, D. M. Codes for sound-source location in non-tonotopic auditory cortex. *J. Neurophysiol.* **80**, 863–881 (1998).
9. Xu, L., Furukawa, S. & Middlebrooks, J. C. Sensitivity to sound-source elevation in non-tonotopic auditory cortex. *J. Neurophysiol.* **80**, 882–894 (1998).
10. Najafi, K., Wise, K. D. & Mochizuki, T. A high-yield IC-compatible multichannel recording array. *IEEE Trans. Electron Devices* **32**, 1206–1211 (1985).
11. Middlebrooks, J. C. & Green, D. M. Directional dependence of interaural envelope delays. *J. Acoust. Soc. Am.* **87**, 2149–2162 (1990).
12. Middlebrooks, J. C. Individual differences in external-ear transfer functions reduced by scaling in frequency. *J. Acoust. Soc. Am.* (in press).

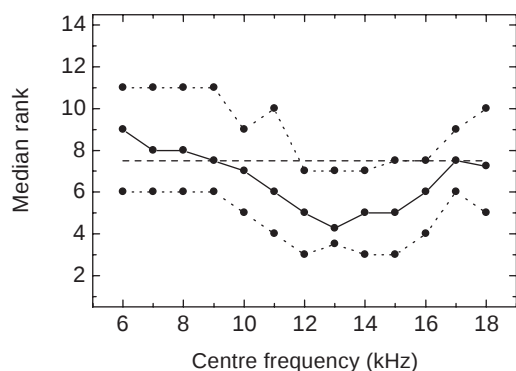


Figure 3 Summary of correspondence between physiological responses and the acoustical model. At each centre frequency, each unit was represented by a median rank as described in the text. Ranks less than 7.5 (dashed line) represent better-than-chance performance. A distribution of 112 to 146 median ranks was compiled at each centre frequency. The solid line represents the medians of the distributions and the dotted lines represent the 25th and 75th percentiles. Data shown are with stimulus levels fixed at 40 dB above neural thresholds. Similar results were obtained in 20- and 30-dB fixed-level conditions and with levels varying in 10-dB steps between 20 and 40 dB above threshold.

Acknowledgements. We thank Z. Onsan for expert technical assistance, and E. Macpherson, B. Mickey, D. Moody, B. Pfingst and J. Schacht for comments on the manuscript. This work was supported by the NIH/NIHCD. The multi-channel recording probes were supplied by the Center for Neural Communication Technology, which is supported by the NIH/NCRR.

Correspondence and requests for materials should be addressed to J.M. (e-mail: jmidd@umich.edu).

The *SIL* gene is required for mouse embryonic axial development and left-right specification

Shai Izraeli*, Linda A. Lowe†, Virginia L. Bertness*, Deborah J. Good‡, David W. Dorward§, Ilan R. Kirsch* & Michael R. Kuehn†

* Genetics Department, Medicine Branch, National Cancer Institute, NIH, Bethesda, Maryland 20889-5105, USA

† Experimental Immunology Branch, National Cancer Institute, NIH, Bethesda, Maryland 20892-1360, USA

‡ Department of Veterinary and Animal Sciences, University of Massachusetts, Amherst, Massachusetts 01003, USA

§ Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, NIH, Hamilton, Montana 59840, USA

The establishment of the main body axis and the determination of left-right asymmetry are fundamental aspects of vertebrate embryonic development. A link between these processes has been revealed by the frequent finding of midline defects in humans with left-right anomalies¹. This association is also seen in a number of mutations in mouse²⁻⁴ and zebrafish^{1,5}, and in experimentally manipulated *Xenopus* embryos⁵. However, the severity of laterality defects accompanying abnormal midline development varies⁶, and the molecular basis for this variation is unknown. Here we show that mouse embryos lacking the early-response gene *SIL* have axial midline defects, a block in midline Sonic hedgehog (*Shh*) signalling and randomized cardiac looping. Comparison with *Shh* mutant embryos⁷, which have axial defects but normal cardiac looping, indicates that the consequences of abnormal midline development for left-right patterning depend on the time of onset, duration and severity of disruption of the normal asymmetric patterns of expression of *nodal*, *lefty-2* and *Pitx2*.

SIL is an early-response gene that was discovered at the site of a leukaemia-associated chromosomal rearrangement⁸. It is ubiquitously expressed in proliferating cells⁹ and during early embryonic development (not shown). To analyse *SIL* function *in vivo*, we deleted exons 3–5, including the translation start sites¹⁰, by homologous recombination in murine embryonic stem cells (not shown). Immunoblot analysis of protein extracts from *SIL*^{-/-} cells confirmed that this is a null allele (not shown). Mice derived from two independent *SIL*^{-/-} embryonic stem-cell clones showed the same phenotype. Heterozygotes are normal, but mutant homozygotes die *in utero* after embryonic day 10.5 (E10.5). Between E7.5 and E8.5, striking developmental anomalies appear in *SIL*^{-/-} embryos. In addition to reduced size and limited developmental progress compared to wild-type embryos (Fig. 1a, b), *SIL* mutants display prominent midline neural-tube defects. These include delay or failure of neural-tube closure and a holoprosencephalic defect, namely the lack of any midline separation at the anterior end of the cranial neural folds (Fig. 1c, d). In addition, left-right development is abnormal. In heterozygous and wild-type embryos the embryonic heart tube always loops to the right (Fig. 1e), whereas in *SIL* mutants the direction of heart looping is randomized (Fig. 1f, g; Table 1).

The transforming growth factor- β (TGF- β) family members

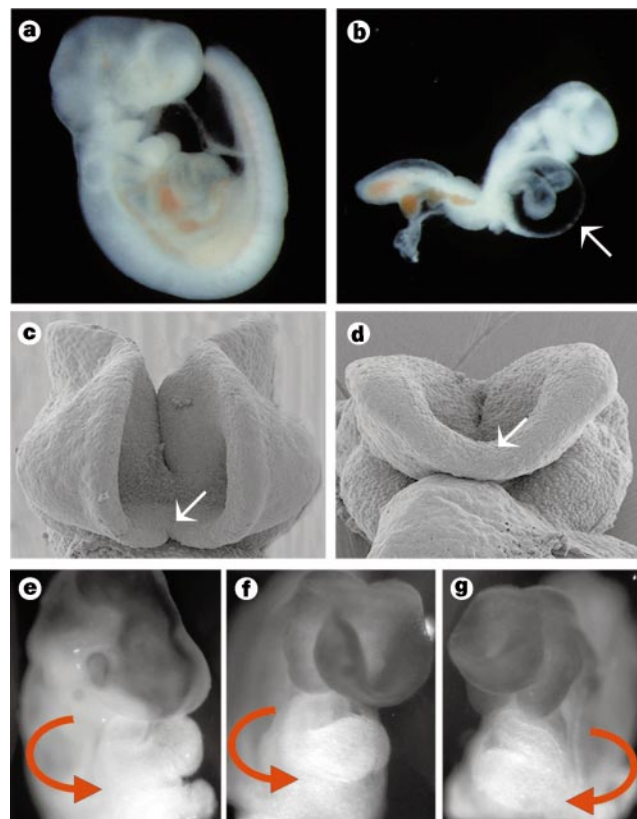


Figure 1 Gross morphology of *SIL*^{-/-} embryos. **a**, Lateral view of normal E9.5 embryo. **b**, E9.5 *SIL* mutant embryo, unturned with abnormal pericardial swelling (arrow). **c**, Scanning electron microscopy (SEM) of normal E8.5 embryo head, frontal view. Arrow, separation at anterior of neural folds. **d**, SEM of E8.5 *SIL*^{-/-} embryo, lacking normal midline separation at the anterior of the neural folds (arrow). **e**, Ventral view of normal E9.5 embryo showing normal rightwards heart looping (direction shown by red arrow). **f**, E9.5 *SIL*^{-/-} embryo with rightwards heart looping. **g**, E9.5 *SIL*^{-/-} embryo with leftwards heart looping.

nodal, *lefty-1* and *lefty-2* and the transcription factor *Pitx2* are all implicated in the establishment of left-right-handed asymmetry in the mouse¹¹⁻¹⁸. *Nodal*, *lefty-2* and *Pitx2* are normally expressed only in the left lateral-plate mesoderm (LPM) before heart looping, with continued expression of *Pitx2* on the left side of the looping heart tube (Fig. 2a–c). In contrast, *SIL* mutants showed bilaterally symmetric expression of *nodal* and *Pitx2* (Fig. 2d, f) at all stages examined (Table 1). For *lefty-2*, most *SIL*^{-/-} embryos also showed bilaterally symmetric expression (Fig. 2e). However, a small number of mutants expressed *lefty-2* only on the right. These were all at very early somite stages, when *lefty-2* expression is first detectable (Table 1), indicating that abnormal *lefty-2* expression on the right side of *SIL* mutants may precede normal left-sided expression. In normal embryos *lefty-1* is expressed on the left side of the prospective floorplate (PFP) at the ventral midline of the neural folds. In *SIL* mutants we found no midline *lefty* expression at any stage examined (Fig. 2e). These results show that *SIL* activity is required to establish the normal pattern of left-right asymmetric gene expression, both in the LPM and in the PFP, from the earliest stages of their expression.

The observed neural-tube defects, including the lack of *lefty-1* expression in the PFP, prompted an analysis of midline developmental pathways. Two key players in midline development are the secreted signalling factor *Shh* and the winged helix transcription factor HNF-3 β (refs 7, 19, 20). Both are normally expressed in the node and notochord at E7.5 to E8.5 (Fig. 3a, g for *Shh*; 3c, e, i, k for HNF-3 β). HNF-3 β is found also in the PFP (Fig. 3i, k), where its expression is dependent on *Shh* signalling from the adjacent node

and/or notochord^{7,21}. In *SIL* mutants, both *Shh* and *HNF-3β* are expressed in the notochord (Fig. 3h for *Shh*; 3j for *HNF-3β*), although by E8.5 the pattern of expression is discontinuous (Fig. 3b for *Shh*) because of gaps in the notochord (not shown). However, there was no *HNF-3β* expression in the PFP at E8.5 (Fig. 3d, j) or at E7.5, when the neural folds and developing notochord appear normal (Fig. 3f, l). This indicates that the lack of neural tube *HNF-3β* expression may be due to defective midline *Shh* signalling, rather than notochord or PFP degeneration. Therefore, we analysed expression of *Patched* and *Gli1*, which are also induced by *Shh* signalling but are more widely expressed in the neural tube than *HNF-3β* (refs 22, 23). Expression of both *Patched* and *Gli1* was greatly reduced in *SIL*^{-/-} embryos (Fig. 3m–p). As with *HNF-3β*, these molecular differences were seen before any gross notochord or neural tube abnormalities (Fig. 3q, r). In contrast, *Smoothed* and *Gli3*, which are not induced by *Shh* signalling but are involved in the *Shh* pathway^{24,25}, showed normal expression (not shown). The markedly reduced expression of multiple *Shh* target genes, despite *Shh* expression in the node and notochord, indicates that *Shh* signalling in the midline may be blocked in *SIL*^{-/-} embryos. A block in *Shh* signalling could explain several similarities

between the *SIL*^{-/-} and mouse *Shh* knockout phenotypes, including the failure of floorplate development, notochord degeneration and holoprosencephaly. However, a pronounced difference is that *Shh*^{-/-} embryos have normal cardiac looping⁷. This finding has been perplexing, because *Shh* has a role in left–right specification in the chick²⁶ and left–right development is highly conserved among vertebrates. To determine whether defects in *Shh* signalling are involved in the abnormal left–right development of *SIL*^{-/-} embryos, we examined heart looping in *Shh* mutants and found it to be normal (Table 1), confirming the original study⁷. Surprisingly, there was bilateral expression of *nodal*, *lefty-2* and *Pitx2* in *Shh*^{-/-} embryos, although only in the anterior right LPM for *lefty-2* and *Pitx2* (Fig. 2g–i). Not all *Shh* mutants had bilateral *lefty-2* expression. Those that did were among the older embryos examined (Table 1), indicating that *lefty-2* turns on in the right LPM after left-sided expression has begun. Midline *lefty-1* expression was not found in any *Shh* mutant (Fig. 2h). Although these results show that *Shh* is involved in regulating left–right gene expression in the mouse, a block in *Shh* signalling alone cannot account for the completely symmetric bilateral pattern of *nodal*, *lefty-2* and *Pitx2* expression found in *SIL* mutants.

In *Xenopus* embryos, removal of the ventral neural tube and notochord before neural-tube closure leads to bilaterally symmetric *nodal* expression and randomized cardiac looping^{5,27}. This indicates that more severe midline defects in *SIL* mutants than in *Shh* mutants might underlie the differences in left–right gene expression and randomized heart looping. Examination of *SIL* mutant embryos by serial sectioning revealed not only abnormal histology but also significantly fewer cells in the neural folds than normal (67 ± 5 cells per section in mutants, 107 ± 5 in heterozygotes at mid-trunk level; $P < 0.0001$, *t*-test; Fig. 4a, b). Because *SIL* is expressed as an immediate early response to growth-factor stimulation⁹, reduced cell numbers might be due to reduced proliferation. However, bromodeoxyuridine (BrdU) labelling was similar in the neural folds of mutant and normal embryos ($68.0 \pm 2.3\%$ BrdU-positive cells in mutants, $63.5 \pm 2.7\%$ in heterozygotes), as was the mitotic index ($1.79 \pm 0.3\%$ in mutants, $1.76 \pm 0.2\%$ in heterozygotes). The lack of a proliferation defect led us to examine programmed cell death as a cause for decreased cell numbers. TUNEL analysis revealed extensive apoptosis (Fig. 4c, d), almost exclusively in the dorsal midline, in the neural folds and somites. To explore further the growth properties of *SIL*^{-/-} cells, we injected *SIL*^{-/-} and *SIL*^{+/-} embryonic stem cells subcutaneously into nude mice. The resulting teratomas were similar in size (1.3 ± 0.4 g, $n = 7$; 1.4 ± 0.6 g,

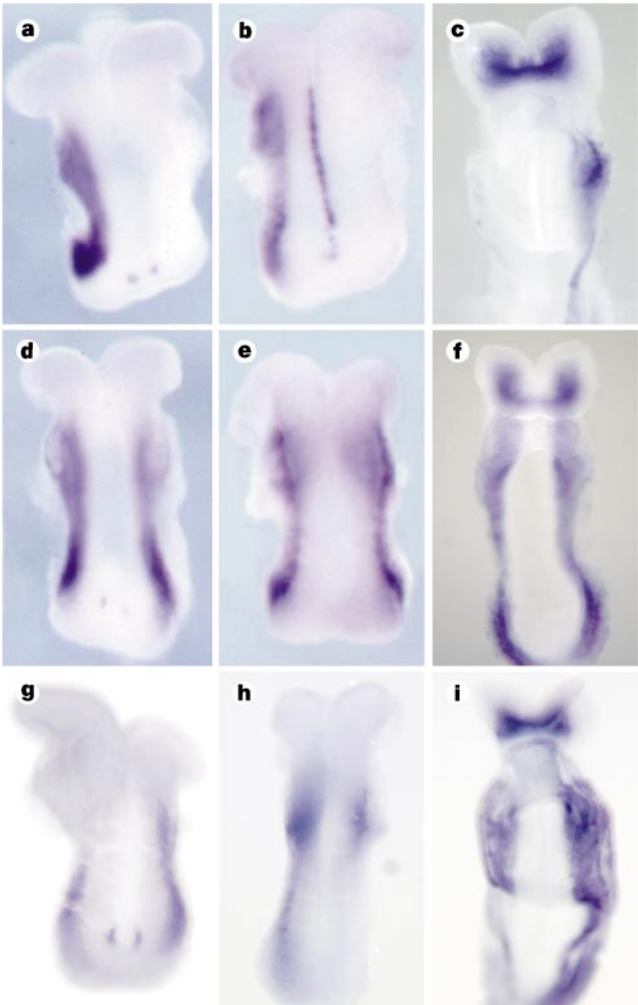


Figure 2 Analysis of left–right asymmetric gene expression at E8.5. **a, b, d, e, g** and **h** show dorsal views; **c, f** and **i** are ventral views. **a**, Normal *nodal* expression in the left LPM. **b**, Normal *lefty-1* expression in the left side of the PFP and *lefty-2* expression in the left LPM. **c**, Normal *Pitx2* expression in head mesenchyme, heart and left LPM. **d–f**, In *SIL*^{-/-} embryos *nodal* (**d**), *lefty* (**e**), and *Pitx2* (**f**) are bilaterally expressed. Note there is no *lefty* expression in the midline. **g–i**, In *Shh* mutant embryos *nodal* (**g**), *lefty* (**h**) and *Pitx2* (**i**) show a different pattern of bilateral expression; again there is no midline *lefty* expression.

Table 1 Left–right development in *SIL* and *Shh* mutants

<i>SIL</i>		Direction of looping of embryonic heart tube				
		Right 34 (46%)	Left 32 (43%)	Ventral 8 (11%)		
		Gene expression pattern in the LPM				
Gene	Somites	Left	Right	Bilateral	Absent	
<i>lefty</i>	3-4		2		3	
	5-6		1	10		
	7-8			3	2	
<i>nodal</i>	3-5			5		
	6-8			6		
<i>Pitx2</i>	6-7			7		
	8-10			6	2	
<i>Shh</i>		Direction of looping of embryonic heart tube				
		Right 29 (97%)	Left 0 (0%)	Ventral 1 (3%)		
		Gene expression pattern in the LPM				
Gene	Somites	Left	Right	Bilateral	Absent	
<i>lefty</i>	3-4	1			1	
	5-6	2		1		
	7-8	1		1		
<i>nodal</i>	3-5			1		
	6-8			5	1	
<i>Pitx2</i>	6-7			4		
	8-10			5		

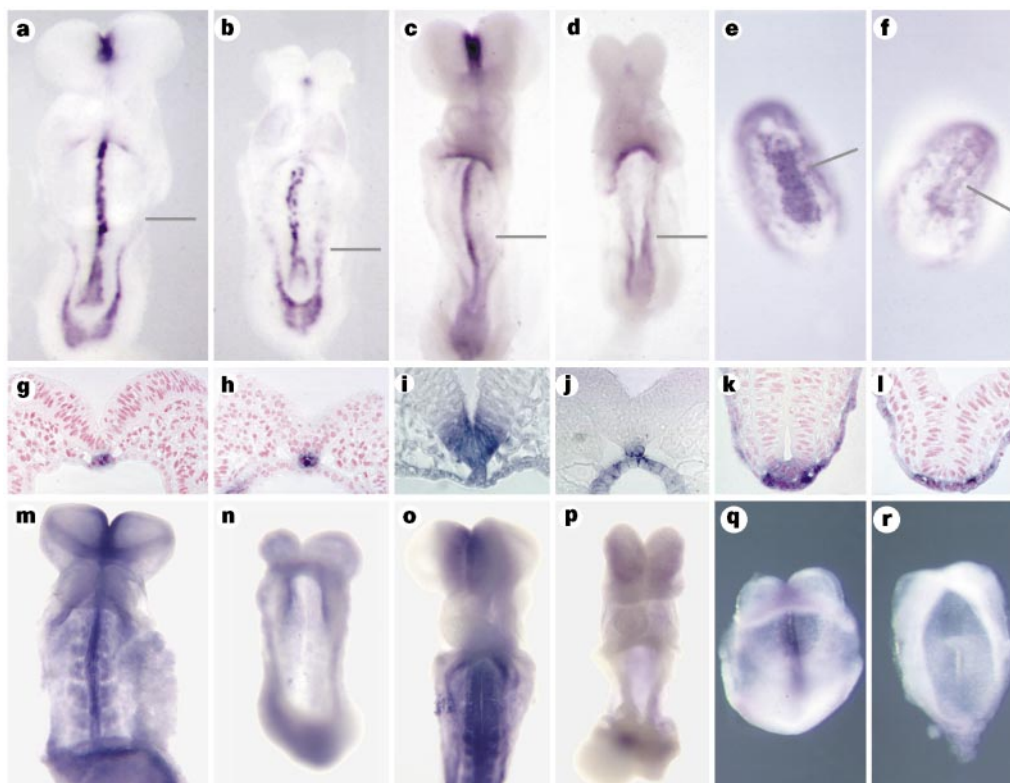


Figure 3 Analysis of Shh signalling. **a–f** and **m–r** show ventral views. Lines in **a–f** indicate approximate level of transverse sections ($10\ \mu\text{m}$) shown directly below in **g–l**. Sections (**g, h, k, l**) were counterstained with nuclear fast red. **a, g**, Normal E8.5 *Shh* expression in the notochord and gut endoderm. **b, h**, E8.5 *SIL* mutant with discontinuous *Shh* expression in the notochord. **c, i**, Normal E8.5 *HNF-3 β* expression in the notochord, PFP and gut endoderm. **e, k**, Normal E7.5 *HNF-3 β* expression in the endoderm, notochordal plate and PFP. **d, j, f, l**, *SIL*^{-/-} embryos

have reduced midline *HNF-3 β* expression due to lack of PFP expression at E8.5 (**d, j**) and at E7.5 (**f, l**). Note the abnormal histology of the mutant neural folds (**j**). **m, q**, Normal *Patched* expression at late E8.5 (**m**) and early E8 (**q**). **n, r**, Reduced *Patched* expression in *SIL*^{-/-} embryos at E8.5 (**n**) and early E8 (**r**). **o**, Late E8.5 normal *Gli1* expression. **p**, Downregulation of *Gli1* expression in a late E8.5 *SIL*^{-/-} embryo.

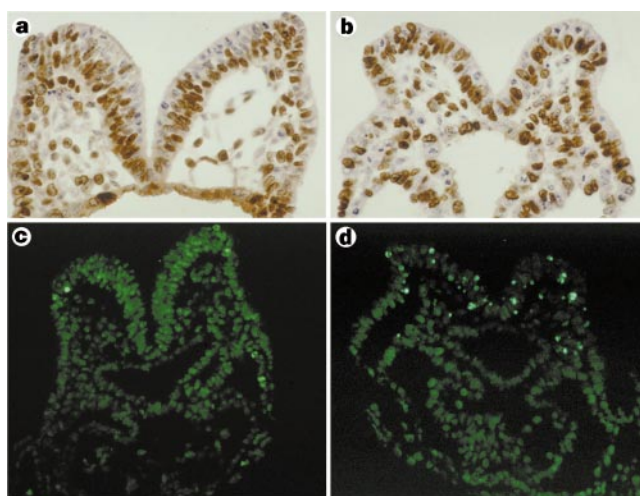


Figure 4 Analysis of cell proliferation and apoptosis at E8.5. **a, b**, Transverse sections ($10\ \mu\text{m}$; haematoxylin counterstained) from mid-trunk level of a normal embryo (**a**) and a *SIL* mutant (**b**) analysed for cell proliferation by BrdU incorporation (brown stain). **c, d**, Analysis of apoptosis by TUNEL assay done *in situ* on sections adjacent to those analysed for BrdU uptake (**c**, normal embryo; **d**, *SIL* mutant). Extensive apoptosis (green fluorescent nuclei) is seen in *SIL*^{-/-} axial and paraxial structures but not in gut or heart.

$n = 10$, respectively) and histology, containing differentiated tissues originating from all three germ layers (data not shown). Therefore, the lack of *SIL* function does not lead to a general growth disadvantage, at least during the chaotic differentiation processes of teratoma formation. Rather, the localized apoptosis in *SIL*^{-/-} embryos indicates that *SIL* is involved in cell survival and growth within the context of the organized developmental program of the midline axial structures.

The early onset of *lefty-2* expression in the right LPM of *SIL* mutants before normal left-sided expression, and the expression of *nodal*, *lefty-2* and *Pitx2* throughout the entire LPM of *SIL* mutants, contrast with the pattern seen in both *Shh* mutants and *lefty-1* mutant embryos¹⁴, in which right-sided expression is delayed and restricted to the anterior LPM. These different patterns point to a role for *SIL* in additional, Shh- and lefty-independent pathways involved in the regulation of *lefty-2* and *Pitx2* expression in the right posterior LPM, from early somite stages. This is consistent with a proposed model in which the midline actively represses the expression of left–right asymmetric genes on the right²⁷, rather than loss of a midline barrier¹⁴. Although *lefty-1* mutants do show laterality defects, most often left pulmonary isomerisms¹⁴, and embryos lacking Shh signalling function also have apparent left isomerisms of the lung^{28,29}, the direction of cardiac looping is normal in *lefty-1* and *Shh* mutants. We propose that the early onset and extended duration of *nodal*, *lefty-2* and *Pitx2* expression throughout the entire right LPM may be the key to why heart looping is randomized in *SIL* mutants but not in *Shh* and *lefty-1* mutants. These observations may also explain why other mouse, human and zebrafish midline mutants show variable co-occurrence and severity of left–right abnormalities. □

Methods

Gene targeting, embryonic stem cells and generation of mice. A 5 kilobase *EcoRI* genomic fragment encompassing exons 3–5 of the mouse *SIL* gene, including putative translation start sites in exons 3 and 4 (ref. 10), was replaced by a PGK-NEO cassette. Gene targeting and generation of chimaeric mice were done by standard methods³⁰. Details of probes and primers for genotyping are available upon request.

Generation of *SIL*^{-/-} cells and teratomas. We generated embryonic stem cells homozygous for the *SIL* null mutation by culturing *SIL*^{+/-} cells in the presence of 4 mg ml⁻¹ G418 for 4 days followed by maintenance in 1 mg ml⁻¹ G418. Teratomas were generated by injecting 2.5 million embryonic stem cells in 0.1 ml PBS subcutaneously into Swiss nude mice.

Cell proliferation and apoptosis studies. E8 pregnant females were injected intraperitoneally with BrdU (100 mg kg⁻¹; Sigma). Embryos were dissected 1 h after injection and fixed in Histochoice (Amresco) for 2–3 h at room temperature, mounted in 2% agarose in PBS to facilitate positioning, embedded in paraffin and cross-sectioned. BrdU was detected by monoclonal antibody (Becton Dickinson) following the manufacturer's protocol. Apoptosis was detected using the *In Situ* Cell Death Detection Kit—Fluorescein (Boehringer Mannheim).

In situ hybridization. Whole-mount *in situ* hybridization, plastic embedding and sectioning were done as described¹¹.

Received 5 February; accepted 5 May 1999.

- Goldstein, A. M., Ticho, B. S. & Fishman, M. C. Patterning the heart's left–right axis: from zebrafish to man. *Dev. Genet.* **22**, 278–287 (1998).
- Oh, S. P. & Li, E. The signaling pathway mediated by the type IIB activin receptor controls axial patterning and lateral asymmetry in the mouse. *Genes Dev.* **11**, 1812–1826 (1997).
- King, T., Beddington, R. S. P. & Brown, N. A. The role of the brachyury gene in heart development and left–right specification in the mouse. *Mech. Dev.* **79**, 29–37 (1998).
- Melloy, P. G. *et al.* No turning, a mouse mutation causing left–right and axial patterning defects. *Dev. Biol.* **193**, 77–89 (1998).
- Danos, M. C. & Yost, H. J. Role of notochord in specification of cardiac left–right orientation in zebrafish and *Xenopus*. *Dev. Biol.* **177**, 96–103 (1996).
- Chen, J. N. *et al.* Left–right pattern of cardiac BMP4 may drive asymmetry of the heart in zebrafish. *Development* **124**, 4373–4382 (1997).
- Chiang, C. *et al.* Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* **383**, 407–413 (1996).
- Aplan, P. D. *et al.* Disruption of the human SCL locus by “illegitimate” V-(D)-J recombinase activity. *Science* **250**, 1426–1429 (1990).
- Izraeli, S. *et al.* Expression of the *SIL* gene is correlated with growth induction and cellular proliferation. *Cell Growth Differ.* **8**, 1171–1179 (1997).
- Collazo-Garcia, N., Scherer, P. & Aplan, P. D. Cloning and characterization of a murine *SIL* gene. *Genomics* **30**, 506–513 (1995).
- Lowe, L. A. *et al.* Conserved left–right asymmetry of nodal expression and alterations in murine situs inversus. *Nature* **381**, 158–161 (1996).
- Collignon, J., Varlet, I. & Robertson, E. J. Relationship between asymmetric nodal expression and the direction of embryonic turning. *Nature* **381**, 155–158 (1996).
- Heymer, J., Kuehn, M. & Ruther, U. The expression pattern of nodal and lefty in the mouse mutant *Ft* suggests a function in the establishment of handedness. *Mech. Dev.* **66**, 5–11 (1997).
- Meno, C. *et al.* Lefty-1 is required for left–right determination as a regulator of lefty-2 and nodal. *Cell* **94**, 287–297 (1998).
- Piedra, M. E., Icardo, J. M., Albajar, M., Rodriguez-Rey, J. C. & Ros, M. A. *Pitx2* participates in the late phase of the pathway controlling left–right asymmetry. *Cell* **94**, 319–324 (1998).
- Ryan, A. K. *et al.* *Pitx2* determines left–right asymmetry of internal organs in vertebrates. *Nature* **394**, 545–551 (1998).
- Yoshioka, H. *et al.* *Pitx2*, a bicoid-type homeobox gene, is involved in a lefty-signaling pathway in determination of left–right asymmetry. *Cell* **94**, 299–305 (1998).
- Campione, M. *et al.* The homeobox gene *Pitx2*: mediator of asymmetric left–right signaling in vertebrate heart and gut looping. *Development* **126**, 1225–1234 (1999).
- Ang, S. L. & Rossant, J. HNF-3 beta is essential for node and notochord formation in mouse development. *Cell* **78**, 561–574 (1994).
- Weinstein, D. C. *et al.* The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. *Cell* **78**, 575–588 (1994).
- Roelink, H. *et al.* Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis. *Cell* **81**, 445–455 (1995).
- Goodrich, L. V., Johnson, R. L., Milenkovic, L., McMahon, J. A. & Scott, M. P. Conservation of the hedgehog/patched signaling pathway from flies to mice: induction of a mouse patched gene by Hedgehog. *Genes Dev.* **10**, 301–312 (1996).
- Hynes, M. *et al.* Control of cell pattern in the neural tube by the zinc finger transcription factor and oncogene *Gli-1*. *Neuron* **19**, 15–26 (1997).
- Akiyama, H. *et al.* Cloning of a mouse smoothened cDNA and expression patterns of hedgehog signalling molecules during chondrogenesis and cartilage differentiation in clonal mouse EC cells, ATDC5. *Biochem. Biophys. Res. Commun.* **235**, 142–147 (1997).
- Ruiz i Altaba, A. Combinatorial Gli gene function in floor plate and neuronal inductions by Sonic hedgehog. *Development* **125**, 2203–2212 (1998).
- Levin, M., Johnson, R. L., Stern, C. D., Kuehn, M. & Tabin, C. A molecular pathway determining left–right asymmetry in chick embryogenesis. *Cell* **82**, 803–814 (1995).
- Lohr, J. L., Danos, M. C. & Yost, H. J. Left–right asymmetry of a nodal-related gene is regulated by dorsoanterior midline structures during *Xenopus* development. *Development* **124**, 1465–1472 (1997).
- Litingtung, Y., Lei, L., Westphal, H. & Chiang, C. Sonic hedgehog is essential to foregut development. *Nature Genet.* **20**, 58–61 (1998).
- Motoyama, J. *et al.* Essential function of Gli2 and Gli3 in the formation of lung, trachea and oesophagus. *Nature Genet.* **20**, 54–57 (1998).

30. Good, D. J. *et al.* Hypogonadism and obesity in mice with a targeted deletion of the *Nhlh2* gene. *Nature Genet.* **15**, 397–401 (1997).

Acknowledgements. We thank H. Westphal and N. Takuma for *Shh* knockout mice; A.-M. Walsh and staff for generation of chimaeras; J. Magrath and R. Mollina for animal care; G. Miller for pathology on teratomas; R. Dryfus for photography; P. Thomas for technical help; P. Aplan, H. Hamada, A. Joyner, C. Chiang, H. Hahn and M. Blum for providing probes; and M. Halpern, H. Arnheiter and B. Casey for comments on the manuscript.

Correspondence should be addressed to M.R.K. (e-mail: mkuehn@box-m.nih.gov). Requests for materials should be addressed to I.R.K. (e-mail: kirschi@exchange.nih.gov).

Regions of variant histone His2AvD required for *Drosophila* development

Michael John Clarkson^{*†}, Julian R. E. Wells^{†‡}, Frank Gibson^{*}, Robert Saint[§] & David John Tremethick^{*}

^{*} The John Curtin School of Medical Research, the Australian National University, PO Box 334, Canberra, Australian Capital Territory 2601, Australia

[†] Department of Biochemistry, and [§] Department of Genetics, University of Adelaide, Adelaide, South Australia 5005, Australia

One way in which a distinct chromosomal domain could be established to carry out a specialized function is by the localized incorporation of specific histone variants into nucleosomes. H2AZ, one such variant of the histone protein H2A, is required for the survival of *Drosophila melanogaster*¹, *Tetrahymena thermophila*² and mice (R. Faast *et al.*, in preparation). To search for the unique features of *Drosophila* H2AZ (His2AvD, also referred to as H2AvD) that are required for its essential function, we have performed amino-acid swap experiments in which residues unique to *Drosophila* His2AvD were replaced with equivalently positioned *Drosophila* H2A.1 residues. Mutated *His2AvD* genes encoding modified versions of this histone were transformed into *Drosophila* and tested for their ability to rescue null-mutant lethality. We show that the unique feature of His2AvD does not reside in its histone fold but in its carboxy-terminal domain. This C-terminal region maps to a short α -helix in H2A that is buried deep inside the nucleosome core.

Figure 1 shows the amino-acid sequence of *Drosophila* H2A.1 (top) and the amino-acid differences found in the *Drosophila* H2AZ protein, His2AvD (bottom). A schematic representation of the secondary structure of H2A is shown below the amino-acid sequence. In addition to the histone fold (the region encompassing α -helices 1 to 3), H2A contains two additional short helices that flank the histone fold (helices α N and α C, respectively)³. Amino-acid differences between His2AvD and H2A span the whole sequence of His2AvD including the N and C termini (excluding the extra-long C-terminal tail of His2AvD, His2AvD has a net gain of five serines and three threonines compared to H2A.1). Clearly, some or all of these amino-acid differences must specify the unique function of His2AvD which cannot be provided by H2A. Therefore, to gain a molecular understanding of these special features that define H2AZ function, we replaced the amino acids unique to His2AvD, by changing the coding sequence, with H2A amino-acid residues.

A deletion in the *Drosophila* *His2AvD* gene (*His2AvD*⁸¹⁰) is homozygous lethal¹. However, the precise developmental step at which lethality occurs was unknown. Here, fly stocks containing the *His2AvD* null allele, *His2AvD*⁸¹⁰, were maintained against the TM6b balancer chromosome which carries the dominant markers *Tubby* and *Humoral*. Non-*Tubby* *His2AvD*⁸¹⁰ homozygotes undergo a protracted third instar and then die without entering pupation.

‡ Deceased

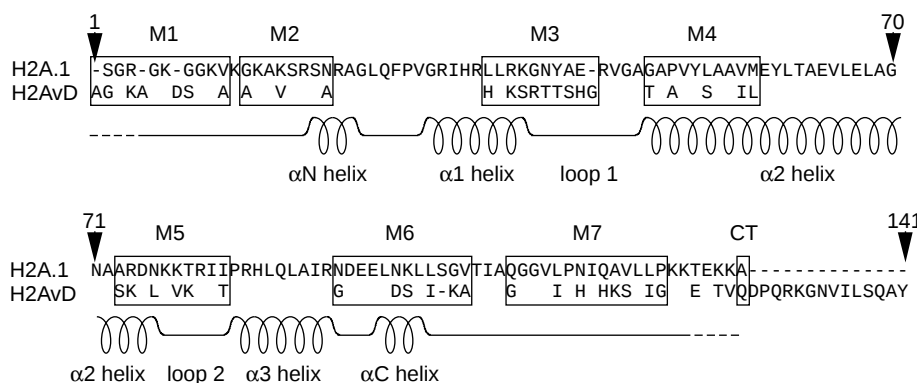


Figure 1 Mutagenesis of the *His2AvD* coding sequence. The amino-acid sequence of *Drosophila* H2A.1 (top) and the amino acids unique to His2AvD (bottom) are shown. Coding sequences on the rescuing *His2AvD* fragment were mutated to encode equivalently positioned H2A.1 amino-acid residues. The

seven transgenes generated encompass all the amino-acid differences between His2AvD and H2A.1. An eighth transgene (CT) was created by removing the terminal 14 amino acids of His2AvD. A schematic representation of the secondary structure of H2A is shown below the sequences.

As *His2AvD*⁸¹⁰ message is maternally transcribed and loaded into the developing oocyte, the arrested development of the *His2AvD* null mutants during the third larval instar indicates that His2AvD derived from the maternal messenger RNA is depleted to subcritical levels at this stage. This null lethality can be rescued with a 4.0 kilobase (kb) genomic fragment containing the *His2AvD* gene. Lines rescued by the wild-type gene generate adult flies that are fertile.

We used site-directed mutagenesis to change the amino-acid sequence of His2AvD. Specifically, 'cassettes' encoding amino-acid residues in the *His2AvD*-rescue DNA fragment were mutated to the equivalent H2A.1 residues. To change all the unique His2AvD amino acids, seven mutants were generated (Fig. 1). To examine the functional contribution of the extended C-terminal domain of His2AvD, an additional mutant was constructed which removed this tail by introducing a stop codon at Gln 127 (CT, Fig. 1). We tested whether these transgenes could rescue null lethality by generating fly lines containing stably integrated wild-type and mutated *His2AvD* transgenes. These lines were then crossed to introduce each transgene into a homozygous *His2AvD*⁸¹⁰ null background. Rescue of the larval lethal phenotype of homozygous *His2AvD*⁸¹⁰ progeny was scored as the appearance of pupae (Fig. 2a) and adults (Fig. 2b) that lacked the dominant TM6b markers *Tubby* and *Humoral*, respectively. For analysis of results, rescue by the wild-type *His2AvD* gene was given a relative value of 100%.

All of the *His2AvD* transgenes containing different H2A-replacement cassettes could rescue null lethality up to pupation with the exception of M6 (Fig. 2a). This non-rescuing transgene contains H2A sequences that lie at the C-terminal end of the protein and not in the histone fold or in the N-terminal tail (Fig. 1). Specifically, transgene M6 has a cassette that encompasses the C-terminal α -helix of H2A that replaces identically located His2AvD residues. As this transgene cannot compensate for His2AvD function, we can conclude that a small C-terminal region of His2AvD, which maps to the extra C-terminal α -helix of H2A and includes only six amino-acid differences (plus one amino-acid deletion) compared to H2A.1 (Fig. 1), is essential for H2AZ function as measured by survival of the fly. This is the first time, to our knowledge, that a small specific region of an individual histone, which lies outside the histone fold, has been shown to be essential for the survival of an organism.

In contrast to the amino-acid segments required to reach the pupal stage, three additional regions of H2AZ (M1, M7 and, to a lesser extent, M4) appear to be important for His2AvD function during the pupal stages (Fig. 2b). Unlike transgene M6, transgenes M1, M4 and M7 do permit some survival to adulthood, but flies transformed with transgene M7 are particularly compromised and on average only about 9% of flies eclose. Cassette 7 lies immediately adjacent to cassette 6 in the C-terminal tail of His2AvD. Therefore, we can conclude that most of the C-terminal region of His2AvD is

important for its function, although the terminal 14 amino acids of His2AvD are not important for survival (Fig. 2b).

Transgene M1 has part of the N-terminal tail of His2AvD swapped with the tail of H2A. This region of His2AvD would be expected to be important for its function because it is the site for post-translational modifications including acetylation and, like the tails of H3 and H4, it may be involved in unique interactions with regulatory proteins⁴. Our results raise the possibility that, during *Drosophila* development, different regions of His2AvD may carry out important functions at different times. However, other explanations exist and further characterization of these mutants during development is required to test this proposal.

To investigate whether the ability of these mutated *His2AvD* genes to rescue null lethality is further compromised at elevated temperatures, we repeated the above experiments at 29 °C instead of 25 °C (Fig. 2c). The overall trend is very similar to the results obtained at non-elevated temperatures, although the rescue ability of some of the transgenes is affected more than others (compare Fig. 2c with Fig. 2b). One marked difference is that transformation with transgene 7 could not rescue null lethality at all. This observation may provide an opportunity to study His2AvD function in the adult fly, and during development, under normal and elevated temperature conditions. A second difference is that the rescue of *His2AvD* null lethality at high temperatures actually increases when the terminal 14 amino acids of His2AvD are removed. In other words, under conditions of moderate heat stress, removing part of the C-terminal tail increases the chance of survival of the adult fly.

Based on the crystal structure of the nucleosome³, we modelled the nucleosome in which normal H2A is replaced with His2AvD (Fig. 3). The C-terminal region of His2AvD critical for its function (cassette 6) is highlighted in red. As this essential region is buried inside the histone core, it is not, at least directly, involved in interactions with DNA. It is more likely that this region is involved in protein-protein interactions, perhaps influencing the stability of the core particle. This seems likely, as analysis of the crystal structure has led to the proposal that the C-terminal α -helix of normal H2A, which interacts with the C-terminal tail of H4, forms a docking domain with α -helix 3 of H2A, and this domain is important for interactions between the H2A-H2B dimer and the H3-H4 tetramer³. (Note that, in the model shown in Fig. 3, the hydrogen-bonding pattern within the docking domain of His2AvD differs significantly from that of H2A). Amino acids 105 to 117, the region of the tail which immediately follows this C-terminal α -helix of H2A (cassette 7), interacts with the N-terminal α -helix of histone H3. This interaction also links H2A with the H3-H4 tetramer³. Therefore, the whole C-terminal region of H2A is involved in nucleosome stability and the amino-acid differences in His2AvD may alter this stability.

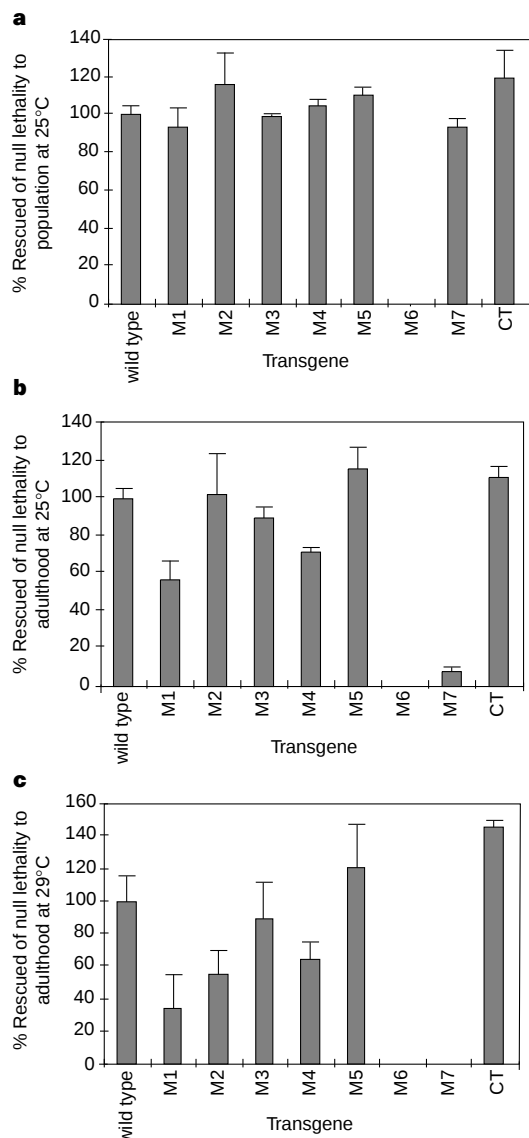


Figure 2 Rescue of *His2AvD*⁸¹⁰ homozygous null lethality by *His2AvD* transgenes. *His2AvD* transgenes (1–7 plus CT) were transformed into *Drosophila* and tested for their ability to rescue *His2AvD*⁸¹⁰ null lethality. The relative rescue provided by each *His2AvD* transgene was determined from three independent transgenic lines and normalized against the wild type (100%). **a**, The percentage of homozygous *His2AvD*⁸¹⁰ individuals reaching pupation. **b**, The percentage of homozygous *His2AvD*⁸¹⁰ individuals reaching adulthood at 25°C. **c**, The percentage rescue of *Drosophila* reaching adulthood at 29°C. The actual level of rescue by the wild-type gene, with 33.3% being the expected value, was 30.8 ± 2.7%, 30.6 ± 1.8% and 23.9 ± 3.5% for **a–c**, respectively.

Any changes in the stability of the core particle would have a direct effect on the ability of remodelling machines⁵ and transcription factors⁶ to disrupt and access chromatin. Such changes in the stability or structure of the nucleosome could also potentially influence the formation of higher-order chromatin structures. For example, the C-terminal α-helix of H2A begins with glutamate, and this residue is part of an acidic patch which, in the nucleosome crystal, interacts with the positively charged N-terminal tail of H4 which protrudes from an adjacent nucleosome core particle³.

In conclusion, our results identify regions that play key roles in the specificity of function of a variant histone and also highlight the multifunctional nature of a nucleosome. Our hypothesis that the incorporation of *His2AvD* alters the stability of the nucleosome

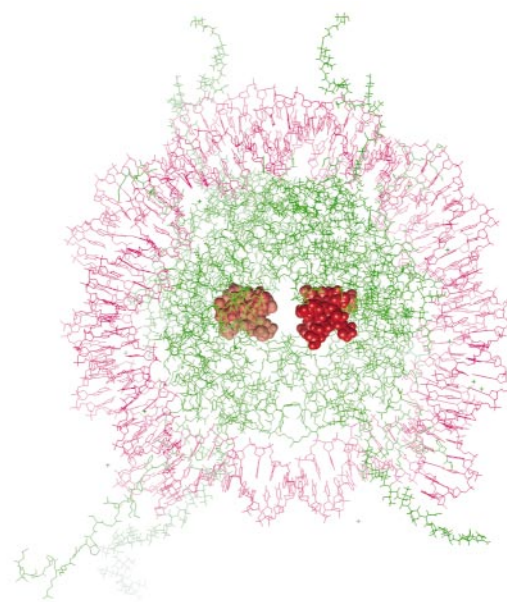


Figure 3 Molecular modelling of the nucleosome core containing *His2AvD*. The C-terminal domain essential for *His2AvD* function is shown in red (cassette 6, amino-acid residues 92–103).

is currently being tested by *in vitro* chromatin reconstitution experiments.

Methods

***His2AvD* mutagenesis.** *His2AvD* coding sequences were mutated using the Altered Sites *in vitro* mutagenesis system from Promega. Restriction endonuclease fragments of the *His2AvD* gene from the plasmid pC4B¹ were subcloned into the Altered Sites vector, pSelect, and mutagenized according to the manufacturers instructions. Mutagenized sequences were subcloned back into a *His2AvD* gene in the P-element transformation construct, pONIX–AvD, which carries the 4.0-kilobase (kb) rescue DNA fragment (data not shown). Correct incorporation of mutations in each of these plasmid constructs was confirmed by sequence analysis. The oligonucleotides used in each mutagenesis were: M1, 5′-CTATTCGAGTCTGGCCGCGGTAAGGGTGGCAAGGTCAAGGCGAAGGCGG-3′; M2, 5′-GGGCAAGGCCAAGGGCAAGGCGAAGAGCCGTTCCAA TCGCGCGGGTCTTC-3′; M3, 5′-CCGACGCGTTCAGCGTAGTTACCCTTA CGTAACAGACGATGGATGCG-3′; M4, 5′-TGTACGCGTCGGAGCCGGAGC ACCCGTGTACTTGGCTGCCGTAATGGAATACC-3′; M5, 5′-CGAGGAATG ATACGTGTCTTTTGTGTCCCTCGCTGCGTT-3′; M6, 5′-GCGATGGTG ACTCCCGAGAGCAGCTTGTTCAGCTCCTCGTTCGGAAT-3′; M7, 5′-TTCTTGGGGAGCAGCACCGCCTGTATGTTCCGAAGACACCGCCTTGAG CG-3′; CT, 5′-GAAGGAGGAAACGGTGTAGGATCCGCAGCGGAAG-3′.

Rescue of *His2AvD* null mutant lethality by *His2AvD* mutants. *Drosophila His2AvD* is encoded by a single copy gene on chromosome three. A 311-base-pair (bp) deletion in this gene which removes the second exon is homozygous lethal (*His2AvD*⁸¹⁰)¹. Flies with *His2AvD* transgenes on the X or second chromosome were crossed to generate stocks homozygous for the transgene and heterozygous for *His2AvD*⁸¹⁰. To test the function of mutant *His2AvD* genes, these stocks were then self-crossed to produce progeny with the *His2AvD* transgene in a homozygous *His2AvD*⁸¹⁰ null mutant background. When normal *His2AvD* function is provided by the transgene, one third of the offspring are homozygous for the null mutation and the remaining two thirds are heterozygous (the *TM6b* balancer chromosome is homozygous lethal). Larvae and flies homozygous for *His2AvD*⁸¹⁰ were scored by the absence of the *Tubby* and *Humoral* markers, respectively, which are carried on the *TM6b* chromosome. DNA was integrated into the germline chromosomes of *Drosophila* using the transposition mechanism of P-elements as described^{7,8}. Transgenic flies were identified by production of eye pigmentation from a *mini-w*⁺ gene present on the transformation construct.

Molecular modelling. The Brookhaven protein data bank file of the *Xenopus*

nucleosome core particle (pdb1aoi.ent) was used. The *Xenopus* H2A peptide served as a template to model the variant *Drosophila* His2AvD peptide using the MSI/Biosym programs InsightII and Homology. To visualize the variant His2AvD peptide in the whole nucleosome structure, a copy of the modelled His2AvD was superimposed on to each of the two copies of the *Xenopus* H2A peptide and the latter were removed.

Received 16 February; accepted 20 April 1999.

1. van Daal, A. & Elgin, S. C. R. A histone variant, H2AvD, is essential in *Drosophila melanogaster*. *Mol. Biol. Cell* **3**, 593–602 (1992).
2. Liu, X., Li, B. & Gorovsky, M. A. Essential and nonessential histone H2A variants in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **16**, 4305–4311 (1996).
3. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
4. Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S. M. & Grunstein, M. Histone H3 and H4N-termini interact with SIR 3 and SIR 4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell* **80**, 583–592 (1995).
5. Peterson, C. L. & Tamkun, J. W. The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem. Sci.* **20**, 143–146 (1995).
6. Ng, K. W., Ridgway, P., Cohen, D. R. & Tremethick, D. J. The binding of a Fos/Jun heterodimer can completely disrupt the structure of a nucleosome. *EMBO J.* **16**, 2072–2085 (1997).
7. Rubin, G. M. & Spradling, A. C. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353 (1982).
8. Spradling, A. C. & Rubin, G. M. Transposition of cloned P-elements into *Drosophila* germ line chromosomes. *Science* **218**, 341–347 (1982).

Acknowledgements. We thank R. King and J. A. Lake for help in constructing mutants M5, M6, M7 and CT.

Correspondence and requests for materials should be addressed to D.J.T. (e-mail: david.trematrick@anu.edu.au) or R.S. (e-mail: rsaint@genetics.adelaide.edu.au).

G-protein-coupled receptor heterodimerization modulates receptor function

Bryen A. Jordan & Lakshmi A. Devi

Department of Pharmacology, New York University School of Medicine, New York, New York 10016, USA

The opioid system modulates several physiological processes, including analgesia, the stress response, the immune response and neuroendocrine function¹. Pharmacological and molecular cloning studies have identified three opioid-receptor types, δ , κ and μ , that mediate these diverse effects^{2,3}. Little is known about the ability of the receptors to interact to form new functional structures, the simplest of which would be a dimer. Structural and biochemical studies show that other G-protein-coupled receptors (GPCRs) interact to form homodimers^{4,5}. Moreover, two non-functional receptors heterodimerize to form a functional receptor, suggesting that dimerization is crucial for receptor function^{6–11}. However, heterodimerization between two fully functional receptors has not been documented. Here we provide biochemical and pharmacological evidence for the heterodimerization of two fully functional opioid receptors, κ and δ . This results in a new receptor that exhibits ligand binding and functional properties that are distinct from those of either receptor. Furthermore, the κ – δ heterodimer synergistically binds highly selective agonists and potentiates signal transduction. Thus, heterodimerization of these GPCRs represents a novel mechanism that modulates their function.

We have previously reported that δ -receptors exist as homodimers and that agonist treatment modulates the level of dimers¹². Using western blotting, we examined lysates from cells expressing κ -receptors tagged with a Flag epitope to see if κ -receptors exist as dimers. We found that most κ -receptors exist as dimers of relative molecular mass (M_r) 130,000, regardless of the presence or absence of a crosslinker (Fig. 1a). The dimers are stable in 10% SDS buffer (not shown); this is unexpected, because δ -receptors show little or no dimeric forms in the absence of the crosslinker; crosslinking is required to stabilize δ -receptor dimers (Fig. 1a). The identity of this

dimer was confirmed by immunoprecipitation experiments using differentially tagged receptors. We found that an antibody to the *myc*-tagged receptor can co-precipitate Flag-tagged receptors from cells expressing both *myc*-tagged and Flag-tagged receptors (Fig. 1b). The dimerization of κ -receptors is not induced by detergents or extraction conditions, as the receptors could be co-precipitated under a variety of conditions, but only from cells co-expressing *myc*- and Flag-tagged receptors (Fig. 1c), and not from a mixture of cells individually expressing the receptors (Fig. 1c). κ -dimers are destabilized in the presence of reducing agents (Fig. 1d), suggesting the involvement of disulphide bonds in receptor dimerization; recently, both the metabotropic glutamate receptor 5 and the calcium-sensing receptor have been shown to dimerize through disulphide bonds^{13,14}. We found that agonist treatment does not induce monomerization of κ -receptor dimers (Fig. 1e), in contrast to δ -receptor dimers, which monomerize in the presence of agonists¹². These results indicate that the properties of κ -receptor dimers and δ -receptor dimers may be different.

We examined the ability of κ -receptors to heterodimerize with δ - or μ -receptors by co-expressing *myc*-tagged κ -receptors with either Flag-tagged δ -receptors or Flag-tagged μ -receptors. Flag-tagged δ -receptors were detected in material immunoprecipitated using antibodies specific for *myc*-tagged κ -receptors (Fig. 2a). In contrast, Flag-tagged μ -receptors could not be detected under similar co-precipitation conditions (Fig. 2a). These results indicate that κ -receptors selectively dimerize with δ - but not with μ -opioid

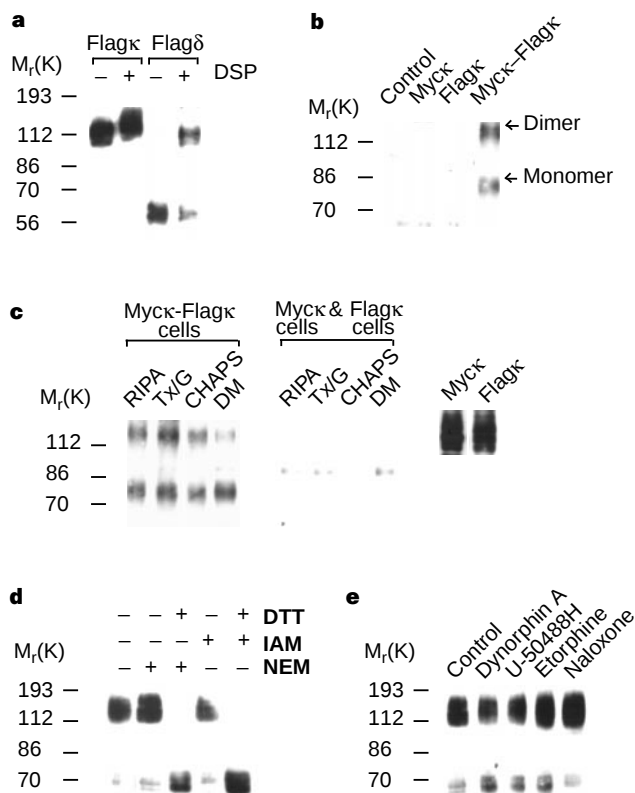


Figure 1 Characteristics of κ -opioid-receptor homodimers. **a**, Immunoblotting of lysates from cells expressing Flag- κ receptors or Flag- δ receptors. **b**, **c**, Myc-tagged κ -receptors can be co-precipitated only from cells expressing both *myc* and Flag-tagged receptors (**b**) under a variety of extraction conditions and not from a mixture of cells individually expressing these receptors (**c**). Expression of *myc*- or Flag-tagged receptors was confirmed by immunoblotting with the appropriate antisera (right panel). **d**, **e**, Treatment of cells expressing κ -receptors with 1 mM DTT for 30 min followed by 5 mM iodacetamide (IAM) or *N*-ethylmaleimide (NEM) results in monomerization (**d**), whereas treatment with 100 nM agonists for 60 min does not (**e**). Immunoblotting experiments used anti-Flag antibodies; immunoprecipitation experiments used anti-*myc* antibodies.

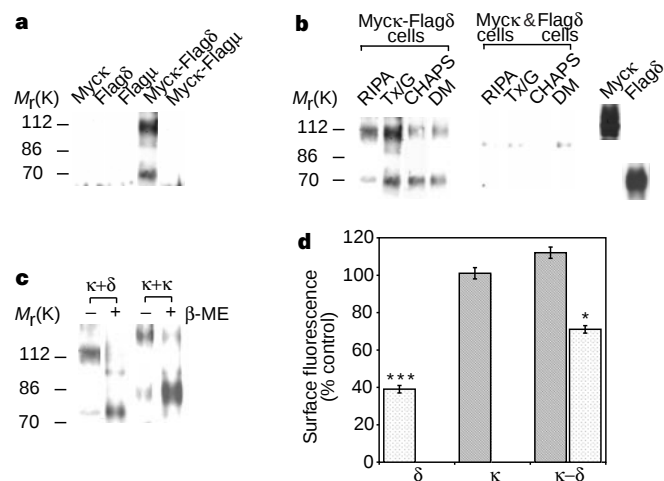


Figure 2 Characterization of κ - δ heterodimers. **a**, κ - δ heterodimers can be immunoprecipitated only from myc- κ - and Flag- δ -expressing cells and not from myc- κ - and Flag- μ -expressing cells. **b**, κ - δ heterodimers can be immunoprecipitated under a variety of extraction conditions and not from a mixture of cells individually expressing these receptors. **c**, Expression of myc- or Flag-tagged receptors in each cell line was confirmed by immunoblotting with the appropriate antisera (right panel). Treatment with β -mercaptoethanol 5% (β -ME) for 5 min results in the destabilization of dimers. **d**, Internalization of receptors in response to 1 μ M etorphine for 60 min. Stippled bars, myc- δ ; shaded bars, Flag- κ . Significant differences from untreated controls are indicated; **P* < 0.05; ****P* < 0.005 (*n* = 3). Immunoblotting experiments used anti-Flag antibodies; immunoprecipitation experiments used anti-myc antibodies.

receptors. κ - δ heterodimers are stable in a variety of detergents and are not induced during solubilization/immunoprecipitation procedures (Fig. 2b). They are also destabilized by a reducing agent (Fig. 2c), suggesting a role for disulphide bonds in κ - δ heterodimerization.

We next examined the effect of heterodimerization on receptor trafficking using cells co-expressing κ - and δ -receptors. Etorphine is a potent, non-selective opioid agonist that binds both δ - and κ -receptors with high affinity. As was shown previously¹⁵⁻¹⁷, etorphine can induce robust internalization of δ - but not κ -receptors in cells individually expressing these receptors (Fig. 2d). In contrast, etorphine cannot induce substantial internalization of δ -receptors in cells expressing both κ - and δ -receptors; the internalization of δ -homodimers in these cells could account for the observed ~25-30% reduction in surface fluorescence (Fig. 2d). These results suggest a role for heterodimerization in altering the trafficking properties of these receptors.

We compared the ligand-binding properties of κ - δ heterodimers with those of κ - or δ -receptors (Table 1), and examined the ability of highly selective agonists^{18,19} and antagonists^{20,21} to compete with [³H]-diprenorphine (a non-selective opioid antagonist) in membranes from cells expressing either κ - or δ - or both κ - and δ -receptors. (Fig. 3a-c). We found that κ -receptors have high affinities for the κ -selective agonist (U69593) and antagonist (norbinaltorphimine). Similarly, δ -receptors have high affinities for the δ -selective agonist ([D-Pen²,D-Pen⁵]enkephalin; DPDPE) and antagonist (TIPP Ψ ; ref. 21). In contrast, κ - δ heterodimers show no significant affinity for either κ - or δ -selective agonists or antagonists (Fig. 3c; Table 1). However, the heterodimer shows a strong affinity for partially selective ligands (Table 1). The properties of these heterodimers are virtually identical to those of the previously reported κ -2-receptor subtype²². Although several studies have reported the

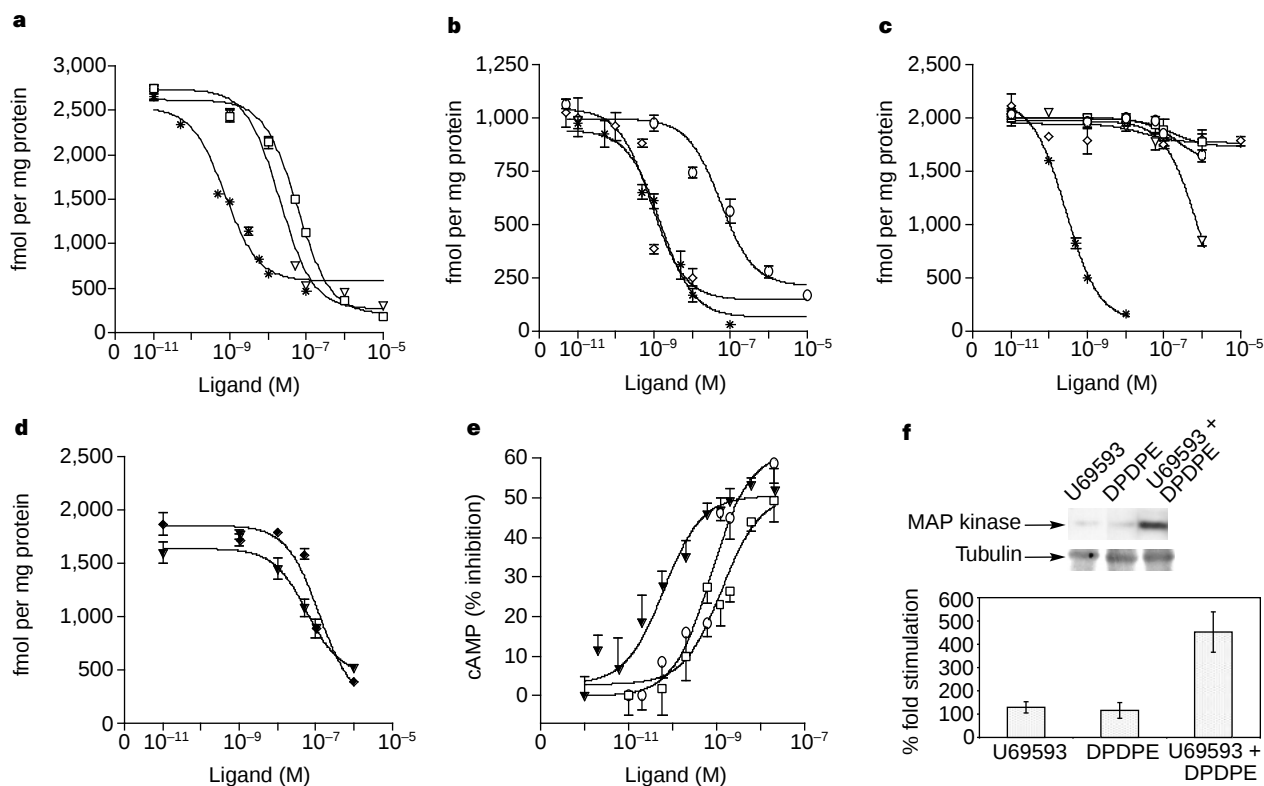


Figure 3 Ligand binding and functional properties. **a-c**, Competition of [³H]-diprenorphine binding by U69593 (square), norbinaltorphimine (triangle), diprenorphine (star), DPDPE (circle) and TIPP Ψ (diamond) in membranes from cells expressing κ - (**a**), δ - (**b**) or κ - and δ - (**c**) receptors. **d**, Displacement of [³H]-diprenorphine by U69593 in the presence of 10 μ M DPDPE (triangle) or DPDPE in the presence of 10 μ M U69593 (diamond). **e**, **f**, Decrease in intracellular cAMP (**e**)

or increase in phospho-MAPK (**f**) by U69593 (square), DPDPE (circle) or U69593 + DPDPE (triangle). In **e**, the 50% inhibitory concentrations (nM) were: U69593, 1.3 $\pm$ 0.7; DPDPE, 0.9 $\pm$ 0.4; U69593 + DPDPE, 0.06 $\pm$ 0.03. Activation of homodimers in these cells could account for the effect seen by individual agonists. Error bars represent s.e.m. (*n* = 3-4).

presence of other subtypes of κ - (ref. 23) and δ - (refs 24, 25) opioid receptors, complementary DNAs corresponding to these subtypes have not been identified despite large-scale efforts by several laboratories. Recent work with δ -receptor knockout mice shows that both the $\delta 1$ and $\delta 2$ receptor subtypes are eliminated in these animals, indicating that the δ -receptor locus may encode both of these subtypes. It is possible that the heterodimerization of δ - or κ -receptors with other GPCRs could form a molecular basis for other receptor subtypes.

We next examined whether the κ - δ heterodimer binds selective agonists synergistically. In the presence of a δ -selective agonist (DPDPE), a κ -agonist (U69593) binds the heterodimer with high affinity (Fig. 3d, Table 1). Similarly, in the presence of the κ -selective agonist (U69593), the δ -agonist (DPDPE) binds with high affinity (Fig. 3d, Table 1). Interestingly, whereas a combination of two selective antagonists also binds with high affinity, a combination of a selective agonist (U69593) and a selective antagonist (TIPP Ψ) does not (Table 1). Also, synergistic binding is not observed in membranes from cells individually expressing κ - or δ -receptors (not shown). Taken together, these results imply that κ - δ heterodimerization results in a new binding site that is able to bind highly selective ligands synergistically.

We examined whether the synergistic binding of agonists leads to increased effector function. The activation of opioid receptors by agonists results in decreased levels of intracellular cyclic AMP and increased levels of phosphorylated mitogen-activated protein kinase (MAPK)²⁶. We found that the potency of individual agonists in reducing intracellular cAMP levels is ~ 10 – 20 -fold lower than that of the two combined (Fig. 3e). Similarly, we found a significant potentiation of MAPK phosphorylation by simultaneous treatment of cells with both agonists as compared to individual agonists (Fig. 3f). These results strongly suggest that the κ - δ heterodimer represents a functional receptor that is activated synergistically by selective ligands.

Our data provide biochemical and functional evidence for opioid-receptor heterodimerization. This is the first direct evidence for the heterodimerization of opioid-receptor types and for two fully functional GPCRs. The heterodimers have greatly reduced affinities for their selective ligands. Interestingly, selective agonists can cooperatively bind to heterodimers and induce synergistic functional responses. Heterodimerization could be a mechanism for activating the receptors on the co-release of selective endogenous peptides. Alternatively, κ - δ -opioid-receptor heterodimers

could represent a hitherto uncharacterized receptor for a specific endogenous opioid peptide. The number of endogenous opioid peptides is far greater than the number of cloned opioid receptors²⁷. Opioid-receptor subtypes resulting from heterodimerization of opioid receptors with other GPCRs could be targets for the action of these endogenous peptides. The physical interactions between GPCRs has enormous ramifications for our understanding of how their actions are regulated. Heterodimerization of opioid receptors also points to additional targets for the development of therapeutic drugs. □

Methods

Generation of cell lines expressing opioid receptors. Cells stably expressing epitope-tagged rat κ - or mouse δ -receptors were generated as described elsewhere²⁸. HEK293 or COS cells co-expressing *myc*-tagged κ - with Flag-tagged κ -, δ - or μ -receptors were generated by transfecting cells using calcium phosphate precipitation¹² and collecting the cells 72 h later for transient expression. For stable expression, CHO cells were transfected with the Flag-tagged κ -receptor cDNA in a geneticin-selectable vector and *myc*-tagged δ -receptor cDNA in a hygromycin-selectable vector (pCEN4, Invitrogen) and selected with $500 \mu\text{g ml}^{-1}$ each of geneticin and hygromycin (Gibco). Surface expression was confirmed by flow cytometry using monoclonal anti-Flag (M1; Sigma) and polyclonal anti-*myc* (c-Myc A14; Santa Cruz) antibodies. Three clones expressing κ - plus δ -receptors at a ratio of approximately 1:1 were used for further studies.

Dimerization, immunoprecipitation and western blotting. Crosslinking with DSP (dithiobis-succinimide propionate; Pierce), SDS-PAGE and western blotting were carried out with lysates of whole cells or membranes essentially as described elsewhere¹², except that cells were lysed in buffers containing the protease-inhibitor cocktail ($10 \mu\text{g ml}^{-1}$ leupeptin, $10 \mu\text{g ml}^{-1}$ aprotinin, 10 mM EDTA, 1 mM EGTA, $10 \mu\text{g ml}^{-1}$ bacitracin, 1 mM pepstatin A, 0.5 mM phenylmethylsulphonyl fluoride and 1 mM E-64) and 100 mM iodoacetamide at 4°C for 60 min (or 23°C when lysing with SDS buffer). In most experiments we used Tx/G buffer (300 mM NaCl, 1% Triton X-100, 10% glycerol, 1.5 mM MgCl_2 and 1 mM CaCl_2 in 50 mM Tris-Cl, pH 7.4) for solubilization. Other solubilization buffers used were RIPA (1% NP-40, 0.5% deoxycholate, 0.5% SDS, 300 mM NaCl), CHAPS (0.5% in 50 mM NaPO_4 buffer, pH 7.4), DM (0.5% dodecyl β -maltoside in 50 mM Tris-Cl, pH 7.4) and SDS (2% SDS in 50 mM Tris-Cl, pH 6.8). For immunoprecipitation, 100–200 μg of proteins were incubated overnight at 4°C with 1–2 μg of polyclonal anti-*myc* antibody. Immunocomplexes were isolated by incubation with 10% v/v of Protein A–Sepharose and analysed by western blotting using monoclonal anti-Flag antibody as described previously¹². κ -receptors are differentially glycosylated in CHO and COS cells; this could account for differences in the size of monomers and dimers. Weak reducing conditions (10 mM dithiothreitol; DTT was added to the immunoprecipitate) were used to eliminate crossreactivity with nonspecific proteins; a portion of dimers are converted to monomers under these conditions.

Binding assays. For membrane preparation, HEK293 cells expressing single or combinations of receptors were washed with PBS buffer, collected with a rubber policeman in 5 mM Tris-Cl buffer, pH 7.4, and incubated for 30 min at room temperature. Cells were disrupted by sonication and subjected to low-speed centrifugation to remove organelles and nuclei. The resulting supernatant was subjected to centrifugation at $50,000g$ for 10 min, and membranes were collected, washed three times, resuspended in 50 mM Tris-Cl, pH 7.4, containing a protease-inhibitor cocktail and stored at -80°C . Membranes were homogenized on thawing and 30–50 μg of membrane proteins were incubated with 0.5 nM or 5 nM ^3H -diprenorphine (40 Ci mmol^{-1} ; NEN/Dupont) for 60 min at 37°C in the absence or presence of 5–8 concentrations of unlabelled ligands; 1 μM unlabelled diprenorphine was used to obtain specific binding. 5 nM ^3H -diprenorphine labels all receptors in κ - δ -expressing cells. Approximately 45% of the specific binding is not displaced by either 10 μM DPDPE or 10 μM U69593; this can be selectively labelled by 0.5 nM ^3H -diprenorphine, so all studies characterizing the κ - δ heterodimer were carried out with 0.5 nM ^3H -diprenorphine. Values for half-maximal inhibitory concentration (IC_{50}) were determined from displacement curves using GraphPad Prism 2.0 and for inhibition constant (K_i) using the Cheng–Prusoff equation²⁹.

Internalization assays. Ligand-induced internalization was carried out by

Table 1 Ligand-binding properties of κ - δ heterodimer

Ligand	K_i (nM)		
	κ	δ	κ - δ
Agonists			
U69593	14.4 ± 0.2	$>1,000$	$>1,000$
DPDPE	$>1,000$	21.8 ± 2.1	$>1,000$
Dynorphin A	1.3 ± 0.41	56.8 ± 0.3	5.6 ± 0.3
EKC	5.7 ± 0.79	105 ± 1.2	2.6 ± 0.2
Etorphine	1.5 ± 0.73	7.9 ± 0.03	7.0 ± 0.3
Bremazocine	0.4 ± 0.11	7.5 ± 0.03	1.2 ± 0.1
Antagonists			
Norbinaltorphimine	4.9 ± 0.3	31.8 ± 1.2	126 ± 9.6
TIPP Ψ	$>1,000$	0.28 ± 0.6	$>1,000$
BNTX	172 ± 19	5.2 ± 0.6	98 ± 19
Naltrindole	44.3 ± 1.1	1.0 ± 0.21	44 ± 1.1
Naloxone	14.9 ± 3.4	43.4 ± 0.3	7.5 ± 0.4
Diprenorphine	0.71 ± 0.27	1.43 ± 1.1	0.3 ± 0.1
Combination of ligands			
U69593 (+10 μM DPDPE)	14.4 ± 0.4	—	9.2 ± 1.4
DPDPE (+10 μM U69593)	—	24.8 ± 1.6	20.0 ± 1.3
Norbinaltorphimine (+10 μM TIPP Ψ)	—	—	0.02 ± 0.003
U69593 (+10 μM TIPP Ψ)	—	—	$>1,000$

Ligand affinities were determined by competition assays using ^3H -diprenorphine as described. Mean $\pm$ s.e.m. ($n = 3$ – 4). BNTX, 7-benzylidenenaltrexone; EKC, ethylketocyclazocine; —, not done.

flow cytometry as described elsewhere¹⁵, except that FITC-conjugated anti-mouse antibody was used to detect the monoclonal anti-Flag antibody (M1; Sigma) bound to Flag- κ receptors and phycoerythrin-conjugated anti-rabbit antibody was used to detect the polyclonal anti-myc antibody (c-Myc A14; Santa Cruz) bound to myc-tagged δ -receptors.

Functional assays. CHO cells co-expressing approximately 1:1 ratio of κ - and δ -receptors were treated with various doses of agonists ($2 \times$ DPDPE, $2 \times$ U69593 or $1 \times$ DPDPE + $1 \times$ U69593) for 5 min at 37 °C. The intracellular cAMP level was measured by radioimmunoassay as described previously²⁸. The level of phosphorylated MAPK was determined by western blotting³⁰. Standardization was with tubulin measured in the same blots using anti-tubulin antibody (Sigma). NIH Image 1.61 software was used to densitize and quantify phospho-MAPK levels. The extent of MAPK phosphorylation in cells treated with 2 nM DPDPE or 2 nM U69593 or with 1 nM DPDPE + 1 nM U69593 is shown in Fig. 3f (upper panel). % fold stimulation refers to the agonist-induced increase in phospho-MAPK levels over untreated levels (taken as control, 100%).

Received 29 March; accepted 6 May 1999.

- Herz, A. *Opioids* Vol. 1 (Springer, Berlin, 1993).
- Miotto, K., Magendzo, K. & Evans, C. J. in *The Pharmacology of Opioid Peptides* (ed. Tseng, L.) 57–71 (Harwood, Singapore, 1995).
- Kieffer, B. L. Recent advances in molecular recognition and signal transduction of active peptides: receptors for opioid peptides. *Cell. Mol. Neurobiol.* **15**, 615–635 (1995).
- Hebert, T. E. & Bouvier, M. Structural and functional aspects of G protein-coupled receptor oligomerization. *Biochem. Cell Biol.* **76**, 1–11 (1998).
- Gouldson, P. R., Snell, C. R., Bywater, R. P., Higgs, C. & Reynolds, C. A. Domain swapping in G-protein coupled receptor dimers. *Protein Eng.* **11**, 1181–1193 (1998).
- Maggio, R., Vogel, Z. & Wess, J. Co-expression studies with mutant muscarinic/adrenergic receptors provide evidence for intermolecular cross-talk between G-protein-linked receptors. *Proc. Natl Acad. Sci. USA* **90**, 3103–3107 (1993).
- Monnot, C. et al. Polar residues in the transmembrane domains of the type I angiotensin II receptor are required for binding and coupling: Reconstitution of the binding site by co-expression of two deficient mutants. *J. Biol. Chem.* **271**, 1507–1513 (1996).
- Jones, K. A. et al. GABA_A receptors function as a heteromeric assembly of the subunits GABA_AR1 and GABA_AR2. *Nature* **396**, 674–679 (1998).
- Kaupmann, K. et al. GABA_A-receptor subtypes assemble into functional heteromeric complexes. *Nature* **396**, 683–687 (1998).
- White, J. H. et al. Heterodimerization is required for the formation of a functional GABA_B receptor. *Nature* **396**, 679–682 (1998).
- Kuner, R. et al. Role of heteromer formation in GABA_B receptor function. *Science* **283**, 74–77 (1999).
- Cvejic, S. & Devi, L. Dimerization of the delta opioid receptor; implications for a function in receptor internalization. *J. Biol. Chem.* **272**, 26959–26964 (1997).
- Romano, C., Yang, W. & O'Malley, K. Metabotropic glutamate receptor 5 is a disulfide-linked dimer. *J. Biol. Chem.* **271**, 28612–28616 (1996).
- Bai, M., Trivedi, S. & Brown, E. M. Dimerization of the extracellular calcium-sensing receptor on the cell surface of CaR-transfected HEK-293 cells. *J. Biol. Chem.* **273**, 23605–23610 (1998).
- Trapaidze, N., Keith, D. E., Cvejic, S., Evans, C. J. & Devi, L. A. Sequestration of the delta opioid receptor: Role of the C terminus in agonist-mediated internalization. *J. Biol. Chem.* **271**, 29279–29285 (1996).
- Keith, D. E. et al. Morphine activates opioid receptors without causing their rapid internalization. *J. Biol. Chem.* **271**, 19021–19024 (1996).
- Chu, P., Murray, S., Lissin, D. & von Zastrow, M. Delta and kappa opioid receptors are differentially regulated by dynamin-dependent endocytosis when activated by the same alkaloid agonist. *J. Biol. Chem.* **272**, 27124–27130 (1997).
- Lahti, R. A., Mickelson, M. M., McCall, J. M. & Von Voigtlander, P. [³H]-U-69593, a highly selective ligand for the opioid κ receptor. *Eur. J. Pharmacol.* **109**, 281–284 (1985).
- Roth, G. et al. [D-Pen², D-Pen⁵]enkephalin analogues with increased affinity and selectivity for δ opioid receptors. *J. Med. Chem.* **258**, 299–303 (1990).
- Portoghesi, P. S., Lipkowski, A. W. & Takemori, A. E. Binaltrophimine and norbinaltrophimine, potent and selective κ -opioid receptor antagonists. *Life Sci.* **40**, 1287–1292 (1987).
- Schiller, P. et al. TIPP Ψ : A highly potent and stable pseudopeptide δ opioid receptor antagonist with extraordinary δ selectivity. *J. Med. Chem.* **36**, 3182–3187 (1993).
- Zukin, R. S., Echbali, M., Olive, D., Unterwald, E. M. & Tempel, A. Characterization and visualization of rat and guinea pig brain δ opioid receptors: Evidence for κ_1 and κ_2 opioid receptors. *Proc. Natl Acad. Sci. USA* **4061**–4065 (1988).
- Nock, B. in *The Pharmacology of Opioid Peptides* (ed. Tseng, L.) 29–56 (Harwood, Singapore, 1995).
- Traynor, J. R. & Elliott, J. Delta opioid receptor subtypes and cross talk with mu receptors. *Trends Pharmacol. Sci.* **14**, 84–86 (1993).
- Zaki, P. A. et al. Opioid receptor types and subtypes: the delta receptor as a model. *Annu. Rev. Pharmacol. Toxicol.* **36**, 379–401 (1996).
- Jordan, B. & Devi, L. A. Molecular mechanisms of opiate receptor signal transduction. *Br. J. Anaesth.* **81**, 12–19 (1998).
- Lord, J. A. H., Waterfield, A. A., Hughes, J. & Kosterlitz, H. W. Endogenous opioid peptides: Multiple agonists and receptors. *Nature* **267**, 495–499 (1977).
- Cvejic, S., Trapaidze, N., Cyr, C. & Devi, L. A. Thr353, located within the COOH-terminal tail of the delta opiate receptor, is involved in receptor down-regulation. *J. Biol. Chem.* **271**, 4073–4076 (1996).
- Cheng, Y. C. & Prusoff, W. H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I₅₀) of an enzyme reaction. *Biochem. Pharmacol.* **22**, 3099–3102 (1973).
- Polakiewicz, R., Schiefel, S. M., Dörner, L. F., Kansra, V. & Comb, M. J. A mitogen-activated protein kinase pathway is required for mu-opioid receptor desensitization. *J. Biol. Chem.* **273**, 12402–12406 (1998).

Acknowledgements. We thank L. Fricker and S. Cvejic for critical reading of the manuscript and P. Schiller for the gift of TIPP Ψ . This work is supported in part by grants from the NIH (NIDA and NINDS).

Correspondence and requests for materials should be addressed to L.A.D. (e-mail: lakshmi.devi@med.nyu.edu)

The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η

Chikahide Masutani*, Rika Kusumoto*†, Ayumi Yamada*‡, Naoshi Dohmae§, Masayuki Yokoi*§, Mayumi Yuasa*, Marito Araki*‡§, Shigenori Iwai||, Koji Takio§ & Fumio Hanaoka*§

* Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565-0871, Japan

† Nara Institute of Science and Technology (NAIST), 8916-5 Takayama, Ikoma, Nara 630-01, Japan

‡ Graduate School of Pharmaceutical Science, Osaka University, 1-6 Yamada-oka, Suita, Osaka 565-0871, Japan

§ Institute of Physical and Chemical Research (RIKEN), Wako-shi, Saitama 351-0198, Japan

|| Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

Xeroderma pigmentosum variant (XP-V) is an inherited disorder which is associated with increased incidence of sunlight-induced skin cancers. Unlike other xeroderma pigmentosum cells (belonging to groups XP-A to XP-G), XP-V cells carry out normal nucleotide-excision repair processes but are defective in their replication of ultraviolet-damaged DNA^{1,2}. It has been suspected for some time that the XPV gene encodes a protein that is involved in trans-lesion DNA synthesis, but the gene product has never been isolated. Using an improved cell-free assay for trans-lesion DNA synthesis, we have recently isolated a DNA polymerase from HeLa cells that continues replication on damaged DNA by bypassing ultraviolet-induced thymine dimers in XP-V cell extracts³. Here we show that this polymerase is a human homologue of the yeast Rad30 protein, recently identified as DNA polymerase η (ref. 4). This polymerase and yeast Rad30 are members of a family of damage-bypass replication proteins^{5–10} which comprises the *Escherichia coli* proteins UmuC and DinB and the yeast Rev1 protein. We found that all XP-V cells examined carry mutations in their DNA polymerase η gene. Recombinant human DNA polymerase η corrects the inability of XP-V cell extracts to carry out DNA replication by bypassing thymine dimers on damaged DNA. Together, these results indicate that DNA polymerase η could be the XPV gene product.

To obtain the complementary DNA coding for the XP-V-correcting DNA polymerase from HeLa cells³, we digested the protein with lysyl endopeptidase to obtain peptides suitable for amino-acid sequence analysis. We isolated four partial amino-acid sequences, from which we designed DNA probes for screening a cDNA library constructed from HeLa poly(A)⁺ RNA. The complete nucleotide sequence was determined of a positive clone with a 3.5-kilobase (kb) insert. The first ATG, preceded by an inframe stop codon, initiates an open reading frame (ORF) encoding a protein of 713 amino acids. The sequence contains all four of the determined amino-acid sequences, indicating that the clone codes for the full-length protein (Fig. 1a). The calculated relative molecular mass (M_r) of the protein, 78.4K, was larger than that of the purified protein from HeLa cells (54K) derived from SDS-PAGE³. A polypeptide corresponding to E⁴⁹⁵-N⁵¹¹ (where notation represents the single-letter amino-acid code and residue position) ended in N, whereas K would be expected from lysyl endopeptidase digestion. Thus, the protein purified from HeLa cells may be truncated from amino acid S⁵¹² to the carboxy terminus, giving a calculated M_r of 55.1K. In support of this idea, mass spectrometry of the E⁴⁹⁵-N⁵¹¹ peptide gave an

M_r of 1.704K, which is close to that calculated from the sequence (1.703K). These results suggest that the polypeptide was from the C terminus of the purified protein, but we do not know whether truncation occurred in the cells or during the course of protein purification. As the HeLa protein was still active, at least in our cell-free system³, the C-terminal region of the protein is probably not responsible for the enzymatic activity.

Searching the various databases for sequence homology to the ORF, we discovered that the N-terminal region of the protein is similar to the *Saccharomyces cerevisiae* proteins Rad30 and Rev1, and to the *E. coli* proteins DinB and UmuC, and to other related proteins. These proteins are all involved in the response to DNA damage, including lesion-bypass DNA synthesis⁵⁻¹⁰. Of these proteins, Rad30 is reported to possess an error-free lesion-bypass DNA-polymerase activity⁴ similar to the one we isolated from HeLa extracts³. Comparison of the amino-acid sequences of these two proteins reveals an overall 19.6% amino-acid identity and 31.9% amino-acid similarity. We conclude that the protein encoded by our cloned cDNA is a human homologue of yeast Rad30. We follow the nomenclature of *S. cerevisiae* and refer to the protein as human DNA polymerase η .

We also found homology between human DNA polymerase η

and other proteins of *Schizosaccharomyces pombe*, *Arabidopsis thaliana* and *Caenorhabditis elegans* in the databases. An alignment of all of these proteins revealed seven conserved domains (regions A–G in Fig. 1b, c). The domains are well conserved in human DNA polymerase η , Rad30, and the putative proteins of *S. pombe*, *A. thaliana* and *C. elegans*, and less conserved in DinB. In UmuC and Rev1, the domains C and G, and C were absent, respectively. Thus, the seven domains may be functionally important elements for these proteins and some of them, such as regions C and G, may be specific for the function of DNA polymerase η . We found no known DNA-polymerase motifs.

Northern blots of RNAs from five XP-V cell lines and from a normal human fibroblast cell line, NB1RGB, were analysed using the isolated cDNA as a probe. As shown in Fig. 2a, four transcripts of different sizes (8.7, 7.5, 4.9 and 2.6 kb) were detected. The intensity of the transcripts in each XP-V cell line varied proportionally, suggesting that these transcripts were from the same gene. In two of the five XP-V cell lines, XP7TA and XP2SA, major differences in band intensity were seen. In contrast, the amount of XPC gene transcript was relatively constant between the cell lines (Fig. 2b). Thus, the principal defects in these XP-V cells are likely to be due to decreased expression of the DNA polymerase η gene. There was a

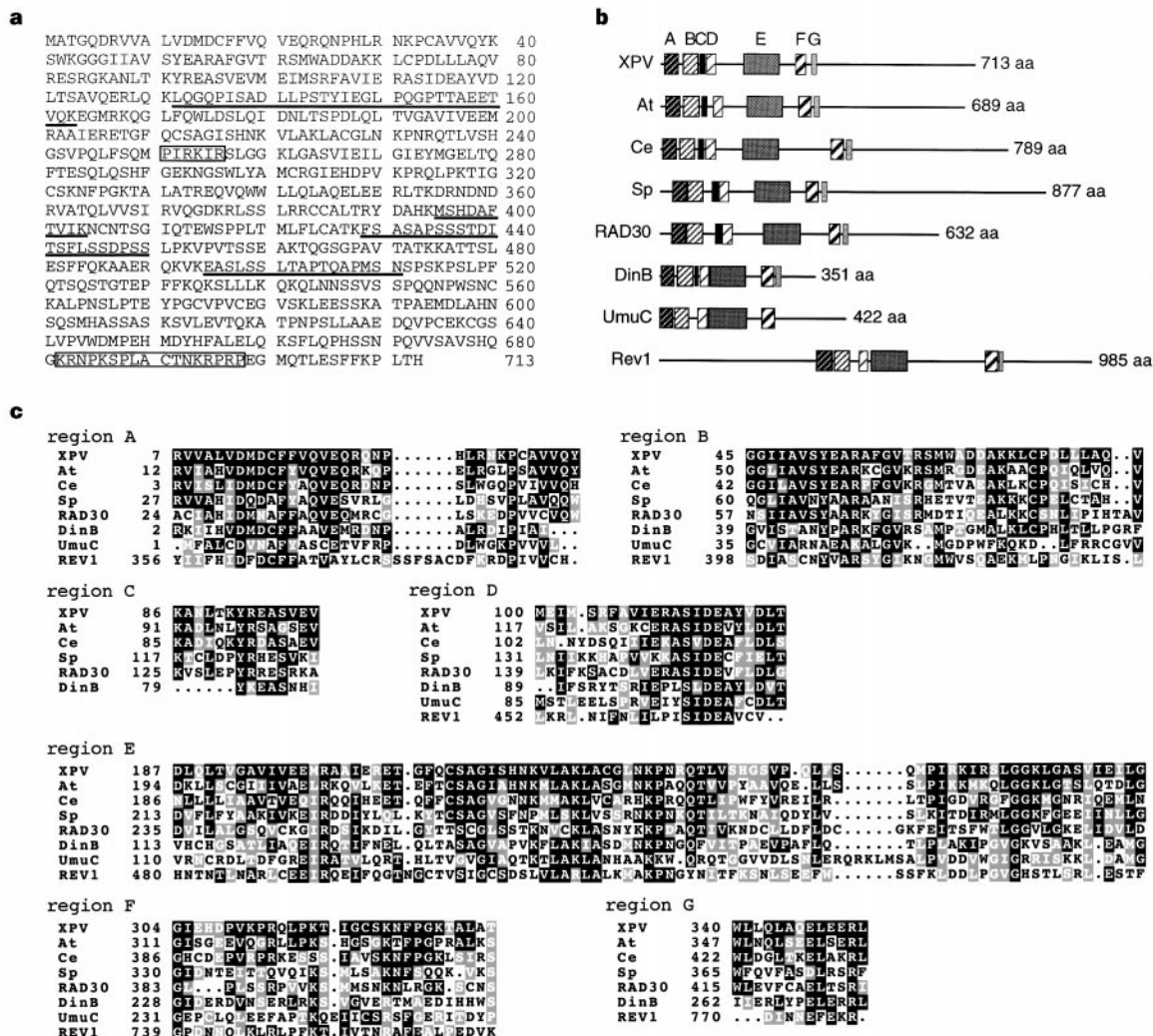


Figure 1 Deduced amino-acid sequence of XPV protein. **a**, Predicted amino-acid sequence of the cloned cDNA. The nucleotide-sequence data reported here will appear in the DDBJ/EMBL/GenBank nucleotide-sequence databases under the accession number AB024313. Amino-acid sequences determined from purified HeLa protein are underlined; two putative nuclear-localization signals found by the PSORT II program are boxed. **b**, Diagram of the conserved regions in human

XPV and related proteins: the conserved regions A–G are indicated. At, *A. thaliana* DinB-like protein T19K24.15; Ce, *C. elegans* UmuC-like protein F53A3.2; Sp, *S. pombe* c16A3.11. **c**, Alignment of the amino-acid sequences of regions A–G. Residues identical and similar to human XPV are indicated in black and grey, respectively.

slight decrease in the amount of transcript in the two cell lines XP1RO and XP30RO, and expression was normal in XP4BE. The transcripts from XP1RO were slightly smaller than those of other cells. To determine the mutation sites in the coding region of the DNA polymerase η gene of XP-V cells, we amplified the entire coding sequence by using the polymerase chain reaction with reverse transcription (RT-PCR) and examined the products by direct sequencing, using the primers shown in Fig. 3a. The mutations found and the predicted proteins encoded by the mutant alleles are summarized in Fig. 3b, c, respectively.

In the cell line XP1RO, the region was deleted between nucleotides 236–376 which contained the start codon. This would explain the smaller transcripts detected by northern-blot analysis. This large deletion probably results from a deletion in the genomic sequence or a mutation in the splice site for mRNA maturation: a definition of the mutation in these cells must await the sequence of the genomic DNA. As for the three cell lines XP30RO, XP4BE and XP7TA, deletions of 13 nucleotides between positions 343–355, four nucleotides at positions 289–292 and two nucleotides at positions 770–771 were found, respectively. These deletions introduce frameshift mutations, which result in truncation of the proteins by new stop codons. Direct sequencing of the amplified cDNA from the cell line XP2SA allowed us to identify a G-to-T mutation at position 1,153, which changes E³⁰⁶ to a stop codon.

As for secondary alleles, we detected a G-to-A mutation at position 1,127 by direct sequencing of the RT-PCR products of XP1RO cDNA amplified with primer S3, which is located in the truncated sequences of the major allele, and primer AS24. The mutation results in the conversion of W²⁹⁷ to a stop codon. As this mutation was not found in the major allele, we conclude that

XP1RO is a heterozygote. However, the normal-sized G1127A allele is only weakly expressed (Fig. 2). For the other XP-V cells that have deletions as major alleles, we tested for products of RT-PCR by using primers specific for the deleted regions of the primary alleles, but were unable to detect any products apart from those expected from the truncated alleles (data not shown). For XP2SA, we subcloned the RT-PCR products and examined the sequences of a random selection of 25 clones, but failed to identify any secondary alleles. In all the XP-V cells examined, we found decreased expression and/or severe mutation of the DNA polymerase η gene, indicating that mutations in this gene are responsible for the defects in XP-V cells, which is consistent with the idea that XP-V consists of one complementation group¹¹.

To determine the activity of the protein encoded by our cloned cDNA, we expressed the recombinant protein, tagged with hexahistidine at its C terminus, in insect cells and purified to near homogeneity (Fig. 4a). The M_r of the purified protein was 83K on SDS-PAGE, which is close to that calculated from the amino-acid sequence. To test the ability of the recombinant protein to correct the defective DNA repair in XP-V cell extracts, we added the protein to damage-bypass DNA-replication assays consisting of template DNA containing the SV40 origin of DNA replication and a single ultraviolet-induced cyclobutane pyrimidine dimer (CPD) located on the leading-strand template³. Extracts from XP-V cells are defective in trans-lesion DNA synthesis^{12–14}, but addition of the recombinant protein significantly increased the amount of replication product in extracts from three XP-V cells (Fig. 4b). Replicative form I (RFI) is the product of complete replication; if the CPD is replicated, RFI will be converted into RFII by treatment with T4 endonuclease V, which introduces a nick at the site of the CPD. The

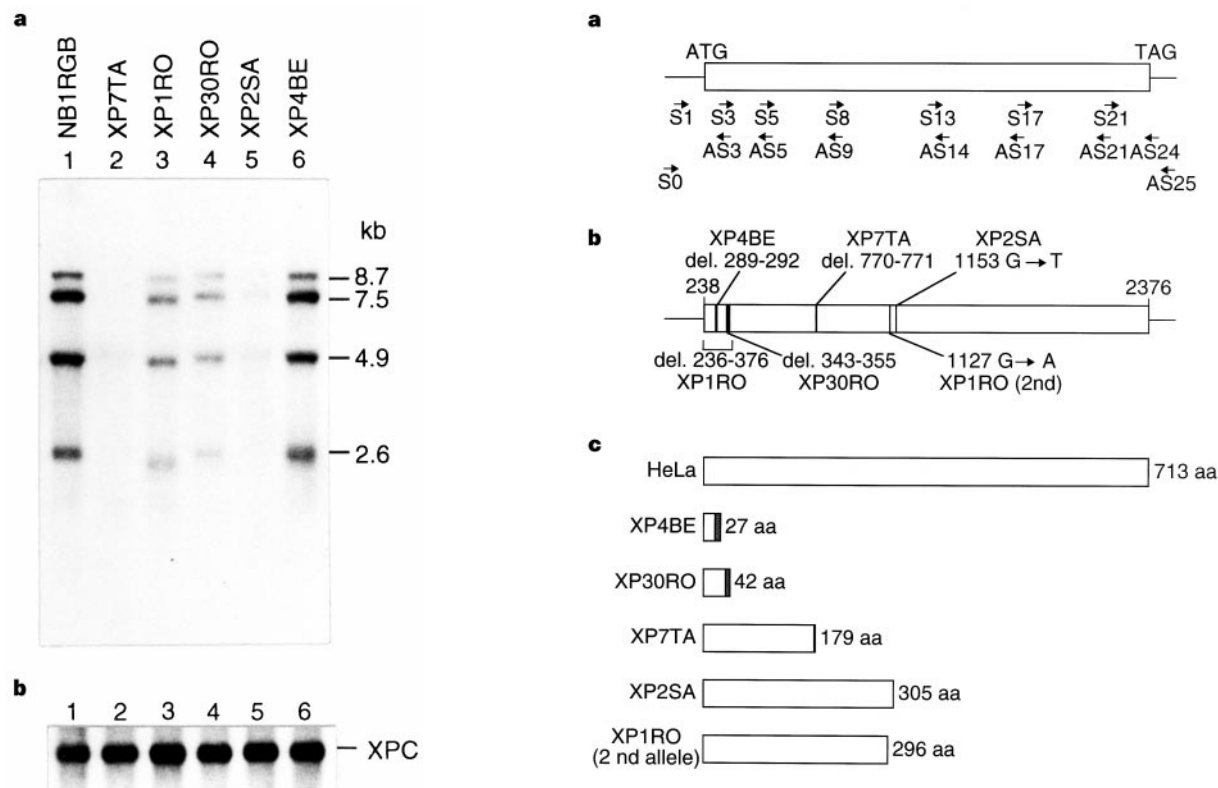


Figure 2 Northern-blot analysis of XPV. **a**, Poly(A)⁺ RNA (2 μ g) prepared from the indicated cells was subjected to agarose gel electrophoresis under denaturing conditions and blotted onto a Hybond-N filter (Amersham Pharmacia). A *KpnI* fragment from the cloned cDNA containing nucleotides 1–2,165 was used as a probe for hybridization. An autoradiograph of the filter is shown. **b**, The same filter was hybridized with the full-length *XPC* cDNA probe²³.

Figure 3 Mutations in the XPV cDNA of XP-V cells. **a**, Primers used for RT-PCR and nucleotide-sequence analyses. The ORF is boxed. **b**, Mutations found in XP-V cells. Mutations in the major alleles of five XP-V cells and in the secondary minor allele of XP1RO are indicated. Numbers indicate nucleotide positions. **c**, Predicted proteins encoded by the mutated genes. The product of the major allele of XP1RO is not represented as it lacks a start codon. Altered coding regions downstream of the mutations, nine, seven, and two amino acids in XP4BE, XP30RO and XP7TA, respectively, are shaded.

increase in the level of replication was sensitive to nicking by T4 endonuclease V, indicating that our protein corrects the defect in trans-lesion DNA synthesis in this cell-free system containing XP-V cell extracts. Transfection of the cDNA into XP7TA cells also corrected their ultraviolet sensitivity (data not shown). DNA

polymerase activity was assayed with a 30-mer template DNA containing a thymine-thymine cyclobutane dimer annealed to a ^{32}P -labelled 13-mer as a primer. As shown in Fig. 4c, the recombinant protein was able to bypass the lesion and synthesize DNA up to the entire 30-mer length, whereas DNA polymerase α stopped just before the lesion and gave 16-mer products. From these results, we conclude that the cDNA coding for the dimer-bypass DNA polymerase is human DNA polymerase η and is the product of the *XPV* gene.

In *S. cerevisiae*, two DNA polymerases, ζ and η , engage in trans-lesion DNA synthesis. DNA polymerase ζ , a complex of the gene products of *Rev3* and *Rev7*, is involved in the error-prone trans-lesion pathway^{15,16}. The human *REV3* homologue has been cloned, indicating that DNA polymerase ζ might be responsible for error-prone trans-lesion DNA synthesis in human cells^{17,18}. In contrast, the yeast *RAD30* gene product, DNA polymerase η , is involved in error-free trans-lesion DNA synthesis^{8,9}, bypassing thymine dimers as effectively as human DNA polymerase η does^{3,4}. The abnormal hypermutability of XP-V cells suggests that the error-free pathway is defective but that the error-prone pathway is still active in these cells¹⁹⁻²¹. Our finding that the *XPV* gene product is an error-free DNA polymerase η accounts well for the phenotype of XP-V cells.

Methods

cDNA cloning. The partially purified dimer-bypass DNA polymerase (15 μg protein) from HeLa cells was subjected to SDS-PAGE and stained with Coomassie blue. The 54K polypeptide was excised from the gel and treated with 0.1 μg *Achromobacter* protease I (ref. 22: a gift from T. Masaki) at 37 °C for 12 h in 0.1 M Tris-HCl (pH 9.0) containing 0.1% SDS and 1 mM EDTA. Peptides generated by protease digestion were extracted from the gel and separated on columns of DEAE-5PW (1 $\times$ 20 mm; Tosoh) and CAPCELL PAK C18 UG120 (1 $\times$ 100 mm; Shiseido) connected in series with a model 1100 (Hewlett Packard) liquid chromatography system; peptides were eluted at a flow rate of 35 $\mu\text{l min}^{-1}$ using a linear gradient of 0–60% of solvent containing 0.075% (v/v) trifluoroacetic acid in 80% (v/v) acetonitrile for 96 min. Selected peptides were subjected to Edman degradation using a model 477A automated protein sequencer connected on-line to a model 120A PTH analyser (Perkin Elmer) and to matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in an instrument (Reflex; Bruker-Franzen Analytik) set to linear mode and using 2-mercaptobenzothiazole as matrix. From the determined amino-acid sequence, we designed primers for PCR with cDNA from HeLa cells: primers 74S2 (ATGWSICAYGAIGCITTYAC) and 75AS2 (TTISWCATIGGIGCYTG) were used for the first PCR with HeLa cDNA, and primers 74S (CAYGAYG-CITTYACNGT) and 75AS (CATIGGIGCYGIGTIGGICNGT) were used for the second PCR, with the first PCR products as templates. After the second PCR, a 0.35-kb product was obtained. Using this fragment as a probe, a $\lambda\text{gt}10$ cDNA library from HeLa cells was screened as described²³. A 3.5-kb insert of the positive clone was obtained by *EcoRI* digestion, subcloned into the *EcoRI* site of pBS-KS(+), and sequenced with an automated DNA-sequence analyser (Amersham Pharmacia ALFred and Applied Biosystems PRISM310).

Mutation analysis. The sequences of the primers were as follows. Sense primers were S0, GATCCCTTCTCGGTTTCTCC; S1, ACTGGACCGCTCC-TAGAAAG; S3, TTGAGGAATAAACCTTGTGCAGTTGTACAG; S5, AGC-CAGTGTGTAAGTGATGG; S8, TTCACACAATAAGGTCCTGGC; S13, AGCTGGTGTGTAGCATTCG; S17, CCATGAGCAATTCACCATCC; and S21, GGATATGCCAGAACACATGG. Antisense primers were AS3, CCACCCCTTC-CATGATTTGTACTGTACAC; AS5, TCCATCACTTCAACACTGGC; AS9, TGCCAGGACCTTATTGTGTGT; AS14, TCCTTGTGAGCATCATAGCGG; AS17, GGATGGTGAATTGCTCATGG; AS21, CCATGTGTTCTGGCATATCC; AS24, ATCCTACAGGCAAGCCTGAG; and AS25, TCCATGCCTGTGAAGA-GATG. First-strand cDNAs of poly(A)⁺ RNAs from HeLa and XP-V cells were obtained using Isogen (Nippon Gene) and Oligotex-dT30 (Takara) and were amplified by PCR using primers S0 and AS25, and ExTaq DNA polymerase (Takara). The products were subjected to a second PCR with primers S1 and AS24. Products of this second PCR were sequenced directly. As for XP1RO, the first-PCR products were amplified with two sets of primers, S1 and AS3, and S3 and AS24, and the resulting products sequenced. As for XP2SA, the second-

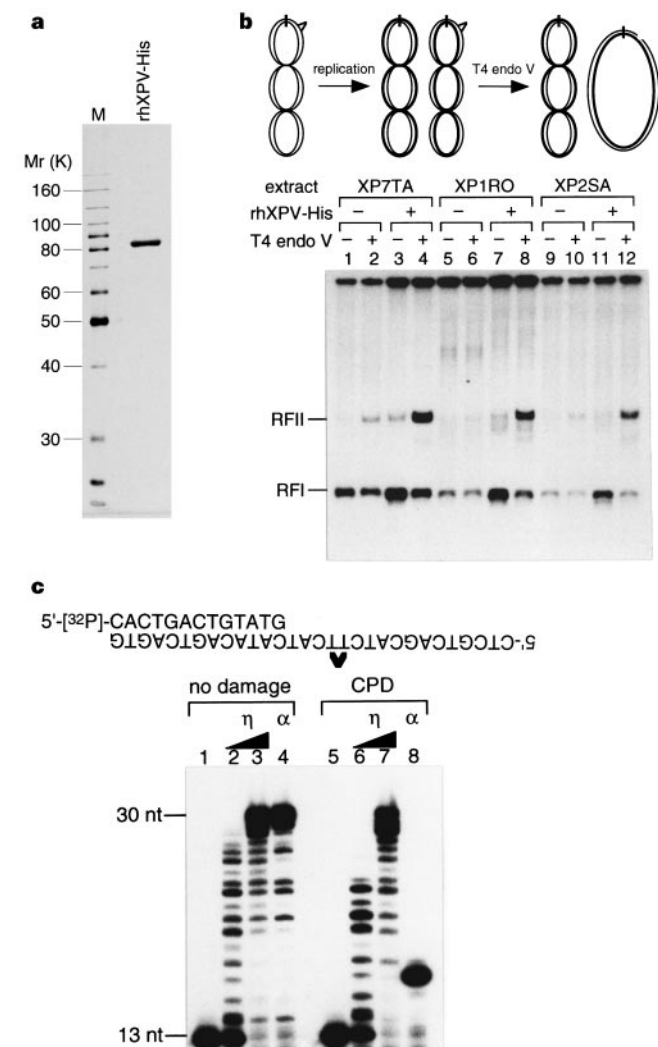


Figure 4 Activity of recombinant XPV protein. **a**, Recombinant XPV protein tagged with hexa-histidine at its C-terminal end was expressed in baculovirus and purified by sequential column chromatography on Hitrap Q, Ni-NTA agarose and MonoS. Detailed procedures will be described elsewhere. A portion (0.3 μg) of the purified fraction was subjected to SDS-PAGE and silver-stained. **b**, Complementation of the defects of XP-V cell extracts. The pBS-SVoriB(CPD), which contains the SV40 origin of DNA replication and a single *cis-syn* thymine dimer on the potential leading strand, was incubated with cell extracts prepared from three XP-V cells³, XP7TA (lanes 1–4), XP1RO (lanes 5–8), and XP2SA (lanes 9–12) in the presence (lanes 3, 4, 7, 8, 11, 12) or absence (lanes 1, 2, 5, 6, 9, 10) of 2 ng recombinant XPV protein. After reaction, DNA products were purified, incubated with (even-number lanes) or without (odd-number lanes) T4 endonuclease V, and fractionated by agarose-gel electrophoresis in the presence of 0.2 $\mu\text{g ml}^{-1}$ ethidium bromide; an autoradiograph of the gel is shown. **c**, DNA polymerase activity. 5' [^{32}P]labelled 13-mer DNA annealed to 30-mer DNA was used as a primed template. The recombinant XPV protein (0.05 ng in lanes 2 and 6, and 0.5 ng in lanes 3 and 7) or DNA polymerase α (0.1 ng; lanes 4 and 8), which was immunoaffinity-purified from mouse FM3A cells, was incubated with primed templates containing (lanes 5–8) or not containing (lanes 1–4) a single *cis-syn* thymine dimer at a site indicated by the bridge. Products were subjected to polyacrylamide (20%) gel electrophoresis under denaturing conditions; an autoradiograph of the gel is shown.

PCR products were cloned into a pGEM-T Easy vector (Promega) and the clones sequenced.

Cell-free assays. XPV activity was examined as follows. Standard reactions of 10 μ l contained 40 mM creatine phosphate-Tris (pH 7.7), 7 mM MgCl₂, 4 mM ATP, 200 μ M each of the other three rNTPs, 20 μ M [α -³²P]dCTP (37 kBq), 100 μ M each of the other three dNTPs, 10 mM dithiothreitol, 250 μ g ml⁻¹ BSA, 100 μ g ml⁻¹ creatine phosphokinase, 2.5 μ g ml⁻¹ pBS-SVoriB(CPD) DNA, 40 μ g ml⁻¹ SV40 large-T antigen, and 5 mg ml⁻¹ proteins from cell extracts³. After incubation at 37 °C for 1 h, reactions were terminated by addition of Na₃EDTA to 25 mM. Mixtures were treated with 25 μ g ml⁻¹ of bovine pancreatic ribonuclease A for 5 min at 37 °C and then with 200 μ g ml⁻¹ of proteinase K for 1 h at 37 °C in the presence of 0.5% SDS. DNA was purified by phenol–chloroform extraction and ethanol precipitation and then dissolved in 30 μ l of buffer V (25 mM Tris-HCl (pH 8.0), 2.5 mM Na₃EDTA, 15 μ g ml⁻¹ BSA). Samples were divided into two and incubated for 0.5 h at 37 °C in the presence or absence of 33 μ g ml⁻¹ of T4 endonuclease V. After addition of one-fifth of the final volume of loading dye (50% glycerol, 0.25% bromophenol blue), samples were subjected to 0.9% agarose-gel electrophoresis with Tris–borate–EDTA buffer containing 0.2 μ g ml⁻¹ ethidium bromide. Autoradiography was done at –80 °C with Fuji New RX X-ray film.

DNA polymerase was assayed as follows. The 5'-[³²P] primer–template DNA was prepared by mixing 13-mer primer, labelled at its 5' end by T4 polynucleotide kinase and [γ -³²P]ATP, and 30-mer DNA, which contained the lesion, at a molar ratio of 1:1. The 30-mer oligomer containing the CPD was chemically synthesized as before²⁴. Standard reactions of 10 μ l contained 40 mM Tris-HCl (pH 8.0), 5 mM MgCl₂, 100 μ M each of the four dNTPs, 10 mM dithiothreitol, 250 μ g ml⁻¹ BSA, 60 mM KCl, 2.5% glycerol, and 4 nM 5'-[³²P] primer–template DNA. After incubation at 37 °C for 15 min, reactions were terminated by addition of 10 μ l formamide and boiling. Products were electrophoresed on a 20% polyacrylamide/7 M urea gel and autoradiographed.

Received 7 April; accepted 25 May 1999.

- Lehmann, A. R. *et al.* Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. *Proc. Natl Acad. Sci. USA* **72**, 219–223 (1975).
- Friedberg, E. C., Walker, G. C. & Siede, W. *DNA Repair and Mutagenesis* (ASM, Washington DC, 1995).
- Masutani, C. *et al.* Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *EMBO J.* (in the press).
- Johnson, R. E., Prakash, S. & Prakash, L. Efficient bypass of a thymine–thymine dimer by yeast DNA polymerase, pol η . *Science* **283**, 1001–1004 (1999).
- Kim, S. R. *et al.* Multiple pathways for SOS-induced mutagenesis in *Escherichia coli*: an overexpression of dinB/dinP results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc. Natl Acad. Sci. USA* **94**, 13792–13797 (1997).
- Smith, B. T. & Walker, G. C. Mutagenesis and more: umuDC and the *Escherichia coli* SOS response. *Genetics* **148**, 1599–1610 (1998).
- Nelson, J. R., Lawrence, C. W. & Hinkle, D. C. Deoxycytidyl transferase activity of yeast REV1 protein. *Nature* **382**, 729–731 (1996).
- McDonald, J. P., Levine, A. S. & Woodgate, R. The *Saccharomyces cerevisiae* RAD30 gene, a homologue of *Escherichia coli* dinB and umuC, is DNA damage inducible and functions in a novel error-free postreplication repair mechanism. *Genetics* **147**, 1557–1568 (1997).
- Roush, A. A. *et al.* Deletion of the *Saccharomyces cerevisiae* gene RAD30 encoding an *Escherichia coli* DinB homolog confers UV radiation sensitivity and altered mutability. *Mol. Gen. Genet.* **257**, 686–692 (1998).
- Walker, G. C. SOS-regulated proteins in translesion DNA synthesis and mutagenesis. *Trends Biochem. Sci.* **20**, 416–420 (1995).
- Jaspers, N. G. J., Jansen-van de Kuilen, G. & Bootsma, D. Complementation analysis of xeroderma pigmentosum variants. *Exp. Cell. Res.* **136**, 81–90 (1981).
- Svoboda, D. L., Briley, L. P. & Vos, J.-M. H. Defective bypass replication of a leading strand cyclobutane thymine dimer in xeroderma pigmentosum variant cell extracts. *Cancer Res.* **58**, 2445–2448 (1998).
- Cordeiro-Stone, M., Zaritskaya, L. S., Price, L. K. & Kaufmann, W. K. Replication fork bypass of a pyrimidine dimer blocking leading strand DNA synthesis. *J. Biol. Chem.* **272**, 13945–13954 (1997).
- Cordonnier, A. M., Lehmann, A. R. & Fuchs, R. P. Impaired translesion synthesis in xeroderma pigmentosum variant extracts. *Mol. Cell. Biol.* **19**, 2206–2211 (1999).
- Nelson, J. R., Lawrence, C. W. & Hinkle, D. C. Thymine–thymine dimer bypass by yeast DNA polymerase ζ . *Science* **272**, 1646–1649 (1996).
- Baynton, K., Bresson-Roy, A. & Fuchs, R. P. Analysis of damage tolerance pathways in *Saccharomyces cerevisiae*: a requirement for Rev3 DNA polymerase in translesion synthesis. *Mol. Cell. Biol.* **18**, 960–966 (1998).
- Xiao, W. *et al.* Identification, chromosomal mapping and tissue-specific expression of hREV3 encoding a putative human DNA polymerase ζ . *Carcinogenesis* **19**, 945–949 (1998).
- Gibbs, P. E. *et al.* A human homolog of the *Saccharomyces cerevisiae* REV3 gene, which encodes the catalytic subunit of DNA polymerase ζ . *Proc. Natl Acad. Sci. USA* **95**, 6876–6880 (1998).
- Wang, Y. C., Maher, V. M. & McCormick, J. J. Xeroderma pigmentosum variant cells are less likely than normal cells to incorporate dAMP opposite photoproducts during replication of UV-irradiated plasmids. *Proc. Natl Acad. Sci. USA* **88**, 7810–7814 (1991).
- Wang, Y.-C., Maher, V. M., Mitchell, D. L. & McCormick, J. J. Evidence from mutation spectra that the UV hypermutability of xeroderma pigmentosum variant cells reflects abnormal, error-prone replication on a template containing photoproducts. *Mol. Cell. Biol.* **13**, 4276–4283 (1993).
- McGregor, W. G., Wei, D., Maher, V. M. & McCormick, J. J. Abnormal, error-prone bypass of photoproducts by xeroderma pigmentosum variant cell extracts results in extreme strand bias for the kinds of mutations induced by UV light. *Mol. Cell. Biol.* **19**, 147–154 (1999).

- Masaki, T., Tanabe, M., Nakamura, K. & Soejima, M. Studies on a new proteolytic enzyme from *Achromobacter lyticus* M497-1. I. Purification and some enzymatic properties. *Biochim. Biophys. Acta* **660**, 44–50 (1981).
- Masutani, C. *et al.* Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and human homolog of yeast RAD23. *EMBO J.* **13**, 1831–1843 (1994).
- Murata, T., Iwai, S. & Ohtsuka, E. Synthesis and characterization of a substrate for T4 endonuclease V containing a phosphorodithioate linkage at the thymine dimer site. *Nucleic Acids Res.* **18**, 7279–7286 (1990).

Acknowledgements. We thank members of F.H.'s laboratory for helpful discussions, K. Tanaka for XP2SA cells, and J. H. J. Hoeijmakers for XP7TA, XP1RO, XP30RO and XP4BE cells. This work was supported by grants from the Ministry of Education, Science, Sports and Culture of Japan, and from the Biodesign Research Program of the Institute of Physical and Chemical Research (RIKEN), M.Y. is a special postdoctoral researcher of RIKEN.

Correspondence and requests for materials should be addressed to F.H. (e-mail: fhanaoka@imcb.osaka-u.ac.jp).

A specific partner for abasic damage in DNA

Tracy J. Matray & Eric T. Kool

Department of Chemistry, University of Rochester, Rochester, New York 14627, USA

In most models of DNA replication, Watson–Crick hydrogen bonding drives the incorporation of nucleotides into the new strand of DNA and maintains the complementarity of bases with the template strand. Studies with nonpolar analogues of thymine and adenine, however, have shown that replication is still efficient in the absence of hydrogen bonds^{1–4}. The replication of base pairs might also be influenced by steric exclusion, whereby inserted nucleotides need to be the correct size and shape to fit the active site against a template base^{5,6}. A simple steric-exclusion model may not require Watson–Crick hydrogen bonding to explain the

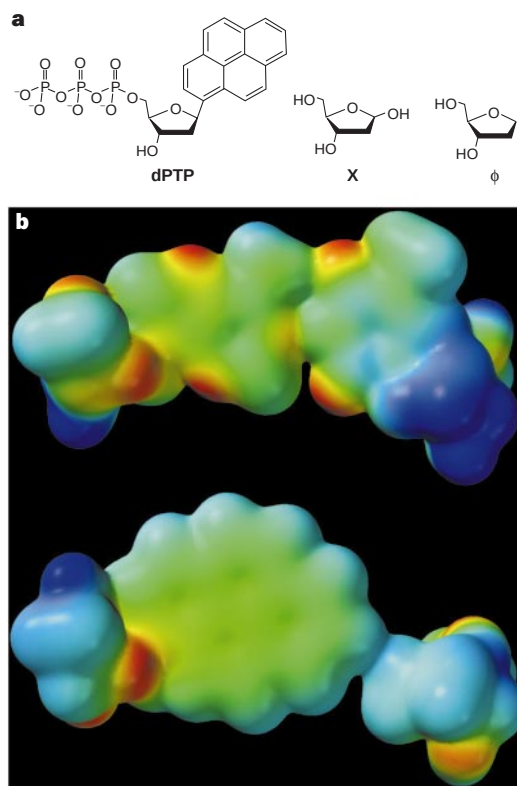


Figure 1 Structures studied. **a**, Structures of dPTP and the abasic nucleosides. **b**, Space-filling models of A-T (top) and P- ϕ (bottom) base pairs in B-form geometry, illustrating the steric fit of pyrene opposite an abasic site.

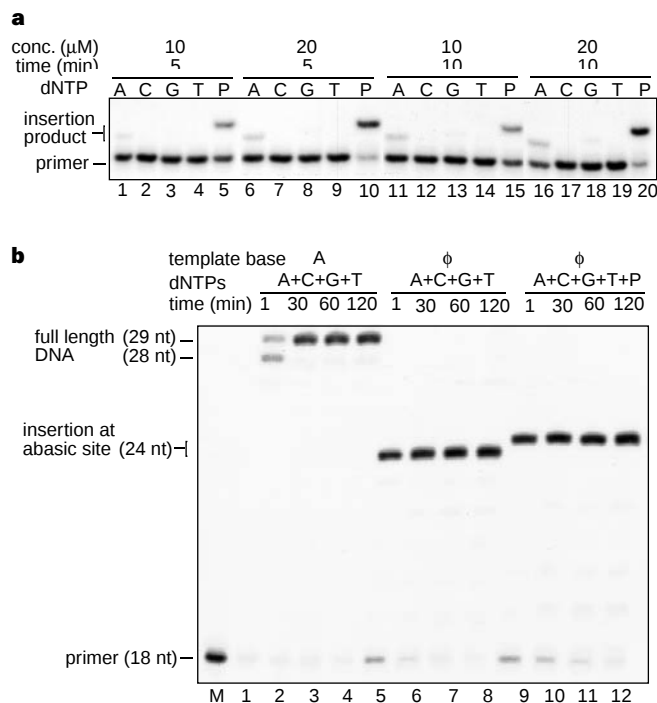


Figure 2 Polymerase insertion of dPTP into templates containing abasic sites. **a**, Autoradiogram showing single-nucleotide insertions on primer-template duplexes containing abasic sites and pyrene nucleoside triphosphate. Sequences of 23-mer primer and 28-mer template are 5'-dTAAACGACTCACTATAGGGAGA and 5'-dACTGXTCTCCCTATAGTGAGTCGTATTA, respectively.

b, Autoradiogram showing running-start multiple nucleotide insertion experiments on an 18/28-nucleotide primer-template duplex, using mixtures of natural nucleotides and dPTP. The 18-mer primer has the sequence 5'-dTAAACGACTCACTATAG.

fidelity of replication, nor should canonical purine and pyrimidine shapes be necessary for enzymatic synthesis of a base pair if each can fit into the DNA double helix without steric strain⁶. Here we test this idea by using a pyrene nucleoside triphosphate (dPTP) in which the fluorescent 'base' is nearly as large as an entire Watson-Crick base pair. We show that the non-hydrogen-bonding dPTP is efficiently and specifically inserted by DNA polymerases opposite sites that lack DNA bases. The efficiency of this process approaches that of a natural base pair and the specificity is 10^2 – 10^4 -fold. We use these properties to sequence abasic lesions in DNA, which are a common form of DNA damage *in vivo*⁷. In addition to their application in identifying such genetic lesions, our results show that neither hydrogen bonds nor purine and pyrimidine structures are required to form a base pair with high efficiency and selectivity. These findings confirm that steric complementarity is an important factor in the fidelity of DNA synthesis.

To test the structural requirements for enzymatic DNA synthesis, we designed a potential non-hydrogen-bonded base pair in which one partner was as small as possible and the other was large enough to occupy the remaining space in the helix. An abasic deoxyribose was selected as the sterically smallest 'nucleoside' available (Fig. 1). We chose a large pyrene deoxynucleoside^{8,9} (dP) as its potential enzymatic pairing partner; models show that its size (surface area, $220 \text{ \AA}^2$) approaches that of a complete thymine-adenine base pair ($269 \text{ \AA}^2$) (Fig. 1b).

The 5'-triphosphate derivative of the pyrene nucleoside (dPTP; Fig. 1) was prepared by standard procedures^{1,10,11} and characterized by ³¹P-NMR and mass spectrometry. To test the ability of the Klenow fragment (Kf) of *Escherichia coli* DNA polymerase I (exo-mutant) to accept this nucleotide as a substrate, we used a primer-template sequence in which the position adjacent to the primer 3' end was either an abasic nucleoside (X) or a tetrahydrofuran abasic analogue (ϕ)^{12–14}. Analytic polyacrylamide gels revealed that sig-

nificant, and apparently efficient, extension of the primer occurred with dPTP opposite the abasic sites at nucleotide concentrations of 10 and 20 μM (Fig. 2a). Moreover, the efficiency was significantly higher than that of dATP opposite abasic sites, which is the basis for the 'A-rule' observed by most polymerases^{14–18}. Insertion of dTTP, dCTP and dGTP were very inefficient. Results were similar with the natural abasic site X (data not shown). To test template selectivity with dPTP, we did similar experiments using varied template bases (A,T,C,G,X, ϕ) with only dPTP added as a free nucleotide (data not shown). The results indicated that dPTP is selectively inserted only opposite abasic sites and that insertion is much less efficient opposite natural DNA bases. We conclude that dPTP is a good polymerase substrate opposite abasic sites and that it is processed with high fidelity.

To test whether dPTP could also be inserted in the presence of mixtures of nucleotides, we did a second experiment in which the primer end was six bases upstream of the abasic site ϕ or X, using the Klenow fragment and the four natural nucleotides (at 800 μM) in the presence or absence of 200 μM dPTP (Fig. 2b). The results were virtually identical for both natural and tetrahydrofuran abasic sites: Fig. 2b shows data for the ϕ analogue. In control experiments using the four natural nucleotides alone, synthesis proceeds rapidly up to and including the abasic site (Fig. 2b, lanes 5–8), with adenine probably being inserted opposite the abasic residue, as noted previously^{14–18}. Further extension of the A-abasic site 'pair' is blocked under these conditions, also as observed previously¹⁸. In the presence of dPTP, synthesis proceeded in a similar way (Fig. 2b, lanes 9–12). However, with dPTP, the stalled band has slower electrophoretic mobility (compare lanes 5–8 and 9–12) owing to the preferential insertion of dPTP opposite ϕ . The same altered mobility was also seen for single-nucleotide insertions of dPTP (Fig. 2b). Thus, dPTP is selectively inserted opposite ϕ , even in the presence of dATP, which is normally considered to be the default choice of the enzyme at this non-instructional lesion. However, the

pyrene nucleotide does not significantly inhibit DNA synthesis for the sequences preceding the abasic site.

As is evident for insertion of dATP opposite abasic sites¹⁸, there is a sudden stall after dPTP insertion (before elongation of the pyrene–abasic pair). The cause of this stalling with either A–abasic or P–abasic pairs is unknown, but the lack of hydrogen-bonding groups in the abasic site may cause unfavourable interaction with polymerase hydrogen-bonding side chains located in the minor groove near the primer terminus^{19–22}. In the case of dPTP insertion, this inhibited elongation was found to be useful in sequencing reactions (see later).

To measure this efficiency of dPTP insertion, steady-state kinetics were followed for single-nucleotide insertions by analysis of the fraction of primer extension at varied nucleotide concentrations^{5,23} (Table 1 and Fig. 3). Of the natural nucleotides, dATP is the most efficiently inserted, as previously reported^{14–18,24}. dPTP is inserted opposite the abasic site very well, with an apparent efficiency ($V_{\max}/K_m$) that is only 3.7-fold lower than insertion of dTTP opposite A in this context. dPTP is inserted with an efficiency that is more than two orders of magnitude greater than that for dATP and 1,500–7,600-fold more than for dGTP, dTTP or dCTP (Fig. 3a). Templates containing natural (X) or analogue (ϕ) abasic sites show little if any difference. These results show that the properties of dPTP make it a better polymerase substrate at abasic sites, thereby breaking the A-rule.

We used kinetics experiments to investigate the template-base preference of dPTP. The results show (Table 1) that dPTP is inserted opposite the abasic site with good template selectivity, despite the lack of hydrogen bonding. The efficiency of insertion opposite ϕ is

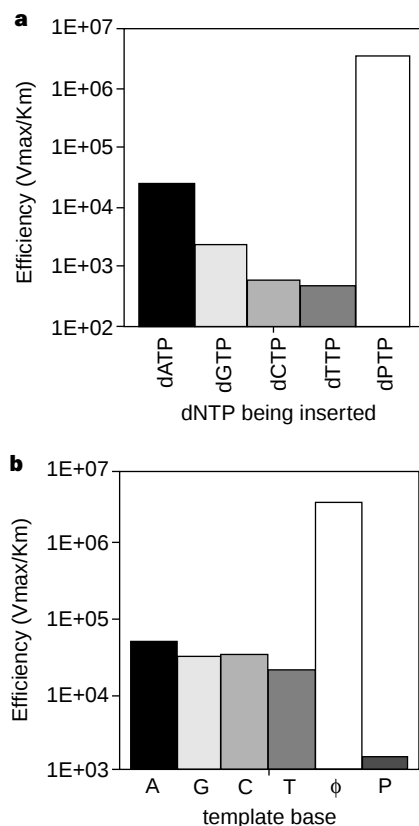


Figure 3 Comparison of the steady-state efficiencies of dPTP insertion. **a**, Fidelity of dPTP insertion opposite an abasic site, relative to insertion of natural nucleotides. **b**, Template selectivity for insertion of dPTP, showing preferential insertion opposite an abasic site rather than opposite a natural base or pyrene itself. Steady-state kinetics for standing-start single nucleotide insertions were carried out as described^{4,6,23}.

two orders of magnitude higher than opposite the natural bases (Fig. 3b). dPTP is inserted opposite an abasic site in preference to another pyrene by a factor of $\sim 10^4$ -fold. We infer that this selectivity of dPTP is a result of good steric accommodation opposite abasic sites (Fig. 1b) but not opposite natural bases. A similar discrimination is shown by dATP, but here the natural nucleotide favours insertion opposite T over abasic sites by 10^2 – 10^3 -fold²⁴ (Table 1). Our results agree with suggestions that hydrogen bonding is not necessary for efficient base-pair synthesis^{1,3,4}. We have previously shown^{1–4} that non-hydrogen-bonding molecules shaped like purines or pyrimidines can be replicated efficiently. The present results demonstrate that this efficiency is almost that of a native base pair, and that purine and pyrimidine structures are not essential for efficient synthesis of a base pair. We conclude that the main requirements for successful nucleotide insertion are that the size and shape of the paired bases are complementary and can be accommodated within the geometry of duplex DNA.

Naturally occurring abasic sites are formed by several mechanisms, including base-excision repair, hydrolysis of alkylated purines, and spontaneous hydrolytic loss of purines^{7,25}. It is important to identify the presence and position of these lesions in order to understand how DNA is damaged and repaired. As abasic sites and alkylated purines are both labile in aqueous solution²⁶, it is difficult to detect abasic sites and distinguish them from other types of DNA damage^{27,28}. Standard polymerase-based sequencing does not distinguish abasic sites from natural bases or from other lesions.

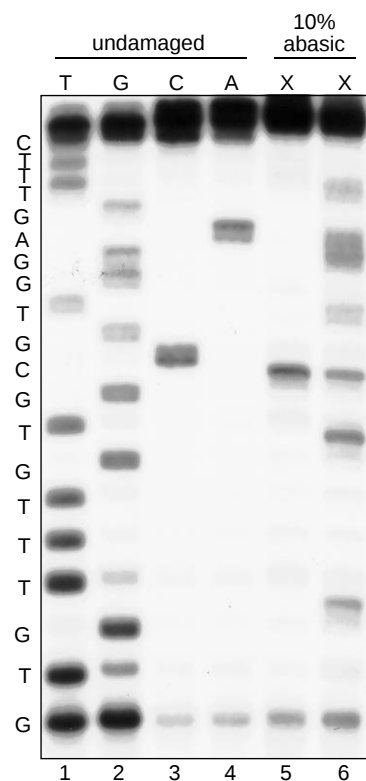


Figure 4 T7 DNA polymerase (exo-)mediated sequencing of a template from the *p53* gene sequence containing single and multiple abasic sites. Lanes 1–4: standard dideoxynucleotide sequencing of undamaged sequence. The full template sequence is 5'-dCTTTGAGGTGCGTGTGTCCTGCTGGGAGAG ACC. Lane 5: template containing a single abasic site (10% abasic, 90% G) in the seventh position; lane 6: template containing 10% abasic damage at each G site in the *p53* template. Reactions in lanes 5 and 6 contain dPTP as well as the four natural nucleotides. Abasic sites were generated by treating DNA containing dU at the desired sites with uracil-DNA glycosidase (UDG). When the precursor template containing U residues rather than abasic sites was used, only a T sequencing ladder was seen (data not shown), with no bands in the abasic lane, which is consistent with the presence of U at these sites.

Table 1 Steady-state kinetic parameters

Nucleoside triphosphate	Template base (X)	$V_{\max}$ (% min ⁻¹)	K_m (μ M)	Efficiency ($V_{\max}/K_m$)	Relative frequency
dPTP	A	0.30(0.03)	6.2(0.4)	4.8×10^4	0.014
	T	0.23(0.03)	11(2)	2.1×10^4	0.0060
	G	0.14(0.02)	4.3(1.2)	3.3×10^4	0.0094
	C	0.12(0.02)	3.5(1.6)	3.4×10^4	0.0097
	X	5.9(1.4)	1.7(0.5)	3.5×10^6	1
	ϕ	9.0(1.4)	3.1(1.2)	2.9×10^6	0.83
	P	0.017(0.008)	11.1(3.6)	1.5×10^3	0.00043
dATP	ϕ	0.93(0.03)	85(10)	1.1×10^4	0.0031
dTTP	ϕ	0.030(0.001)	151(18)	2.0×10^2	0.000057
dGTP	ϕ	0.13(0.007)	114(16)	1.1×10^3	0.00031
dCTP	ϕ	0.027(0.002)	90(9)	3.0×10^2	0.000086
dATP	X	1.80(0.06)	71(10)	2.5×10^4	0.0071
dTTP	X	0.082(0.007)	180(11)	4.6×10^2	0.00013
dGTP	X	0.27(0.03)	114(18)	2.3×10^3	0.00066
dCTP	X	0.060(0.009)	105(10)	5.7×10^2	0.00016
dTTP	A	17.5(3.1)	1.3(0.3)	1.3×10^7	3.7

Steady-state kinetic parameters are shown for insertion of dPTP opposite natural bases, natural abasic sites (X) and a tetrahydrofuran analogue (ϕ). Data were obtained with the Klenow fragment of *E. coli* DNA polymerase I. Values are normalized to a 5 μ M primer-template duplex concentration. Standard deviations are given in parentheses.

Our results suggest that dPTP could be used in the same way as dideoxynucleotide terminators are in sequencing, by acting as a specific marker for abasic sites.

We therefore synthesized a new primer/template combination carrying abasic sites at predetermined positions in different sequence. The 43-mer template was taken from the human *p53* gene near sites known to undergo G $\rightarrow$ T transversion mutation, which could arise from abasic sites after alkylation damage. We included abasic lesions at and near a known mutational hotspot in exon 7 of this gene²⁹. Two separate templates were prepared: one with 10% abasic damage at G at the mutational hotspot, and one with 10% abasic lesions at all sites occupied by G in the normal *p53* sequence. This latter template contains abasic lesions in four different sequence contexts.

In Fig. 4, the dPTP (abasic sequencing) lane shows specific bands at the 10% abasic positions. Dilution of the abasic damaged DNAs with undamaged sequences revealed that the dPTP-induced band could be visualized at a level of 1% abasic damage (data not shown); misinsertion of dPTP opposite natural bases appeared to be limiting below this. Therefore, at least two different polymerases can insert dPTP opposite abasic sites with high efficiency and selectivity, so it may be feasible to analyse naturally derived DNAs for abasic damage by use of cycle sequencing. As pyrene nucleoside is an efficient fluorophore³⁰, it might be useful as a fluorescent marker for this form of DNA damage. □

Methods

Preparation of dPTP. Pyrene β -deoxynucleoside^{6,8} (20 mg, 0.063 mmol) and Proton Sponge (1,8-bis(dimethylamino)naphthalene) (20 mg, 0.10 mmol) were dissolved in trimethylphosphate (0.5 ml) and cooled to 0°C. After dropwise addition of phosphorus oxychloride (7.0 ml, 0.076 mmol), the solution was stirred for 2 h at 0°C. Tributylamine (100 ml, 0.42 mmol) and tributylammonium pyrophosphate (58 mg) in dry dimethyl formamide (0.7 ml) were added and the solution was stirred for 1 min before being quenched by addition of 1 M triethylammonium bicarbonate (10 ml). After stirring for 20 min at room temperature, the reaction was concentrated *in vacuo*. The crude triphosphate was purified by anion-exchange chromatography on a DEAE-cellulose column at 4°C with a 0.1–1.0 M gradient of triethylammonium bicarbonate (pH 7.5). The appropriate fractions were converted to the sodium salt using standard procedures¹¹. The concentration of dPTP was calculated using an extinction coefficient ($33,390 \text{ M}^{-1} \text{ cm}^{-1}$) determined for 1-pyrenemethylamine hydrochloride (at 276 nm) in water. The phosphorus NMR spectrum (D_2O) containing Tris-HCl (pH 7.5, 50 mM) and EDTA (2 mM) gave the three expected triphosphate signals: a doublet at 10 p.p.m. (α -phosphate), a doublet at 13 p.p.m. (γ -phosphate), and a triplet at 24 p.p.m. (β -phosphate). Electrospray mass spectrometry analysis of dPTP

gave the expected mass (found, 579.2 ($\text{M} + \text{Na}^+$); calculated for parent ion, 557.31).

Polymerase reactions. Primers were 5'-end-labelled with [γ -³²P]ATP and T4 polynucleotide kinase. Single nucleotide insertions were made by mixing unlabelled primer and template (final, 10 μ M each) with a buffer (10 $\times$) consisting of Tris-HCl (pH 7.5, 500 mM), MgCl₂ (100 mM), dithiothreitol (10 mM) and acetylated BSA (0.5 mg ml⁻¹). An aliquot of labelled primer (25 pmol) was added, after which the solution was heated at 95°C for 5 min and then allowed to cool to room temperature for over 1 h. Following addition of the Klenow fragment of *E. coli* DNA polymerase I (*exo*-mutant; Amersham; 100 units), the solution was incubated for 2 min at 37°C. Reactions were initiated by adding the DNA/enzyme cocktail (5 μ l) to an equal volume of a 2 $\times$ dNTP stock solution containing Tris (pH 7.5, 200 mM), MgCl₂ (20 mM), mercaptoethanol (6 mM) and dNTP (4.0 or 40 mM). Extension reactions were run for 5 or 10 min, and then quenched by addition of 10 μ l of loading buffer (80% formamide, 1 $\times$ TBE (Tris-borate (90 mM, pH 8.3), EDTA (2 mM)), 0.05% xylene cyanol and bromophenol blue). Running start experiments were conducted under similar conditions but used unlabelled primer/template at 200 nM final concentration, dNTPs at 750 μ M final concentration each (200 μ M dPTP), and were quenched at the indicated times. Reaction products were resolved from unreacted primer by electrophoresis on 20% denaturing polyacrylamide gels and visualized by autoradiography.

Steady-state kinetics. Steady-state kinetics were followed as described^{3,5,23} using the Klenow fragment of *E. coli* DNA polymerase I. Insertion reactions were carried out in 10 μ l in the presence of Tris-HCl (pH 7.5, 50 mM), MgCl₂ (10 mM), dithiothreitol (1 mM) and acetylated BSA (0.05 mg ml⁻¹). Duplex concentration, enzyme concentration, and reaction time were adjusted for different insertion reactions to allow 1–20% primer extension.

Sequencing reactions. Sequencing was done with a slight modification to the procedures described in the Sequenase v2.0 kit (Amersham). For reactions sequencing abasic sites, the reaction contained dNTPs (1.5 mM), dPTP (400 μ M), and NaCl (50 mM), and labelled primer (5'-dGGTCTCTCCAGG ACAGGC; 5 pmol).

Received 16 February; accepted 16 April 1999.

- Moran, S., Ren, R. X.-F. & Kool, E. T. A thymidine triphosphate shape mimic lacking Watson-Crick pairing ability is replicated with high specificity. *Proc. Natl Acad. Sci. USA* **94**, 10506–10511 (1997).
- Goodman, M. F. Hydrogen bonding revisited: geometric selection as a principal determinant of DNA replication fidelity. *Proc. Natl Acad. Sci. USA* **94**, 10493–10495 (1997).
- Morales, J. C. & Kool, E. T. Efficient replication of a DNA base pair between non-hydrogen-bonded nucleoside analogues. *Nature Struct. Biol.* **5**, 950–954 (1998).
- Moran, S., Ren, R. X.-F., Rumney, S. & Kool, E. T. Difluorotoluene, a nonpolar isostere of thymine, codes specifically and efficiently for adenine in DNA replication. *J. Am. Chem. Soc.* **119**, 2056–2057 (1997).
- Goodman, M. F., Creighton, S., Bloom, L. B. & Petruska, J. Biochemical basis of DNA replication fidelity. *Crit. Rev. Biochem. Mol. Biol.* **28**, 83–126 (1993).
- Kool, E. T. Replication of non-hydrogen bonded bases by DNA polymerases: a mechanism for steric matching. *Biopolymers (Nucleic Acid Sciences)* **48**, 3–17 (1998).
- Loeb, L. A. & Preston, B. D. Mutagenesis by apurinic/aprimidinic sites. *Annu. Rev. Genet.* **20**, 201–230 (1986).
- Ren, R. X.-F., Chaudhuri, N. C., Paris, P. L., Rumney, S. & Kool, E. T. Naphthalene, phenanthrene, and pyrene as DNA base analogues: synthesis, structure, and fluorescence in DNA. *J. Am. Chem. Soc.* **118**, 7671–7678 (1996).
- Matray, T. J. & Kool, E. T. Selective and stable DNA base pairing without hydrogen bonds. *J. Am. Chem. Soc.* **120**, 6191–6192 (1998).
- Mishra, M. C. & Broom, A. D. A novel synthesis of nucleoside 5' triphosphates. *J. Chem. Soc., Chem. Commun.* 1276–1277 (1991).
- Hoard, D. E. & Ott, D. G. Conversion of mono- and oligodeoxyribonucleotides to 5'-triphosphates. *J. Am. Chem. Soc.* **87**, 1785–1788 (1965).
- Millican, T. A. et al. Synthesis and biophysical studies of short oligodeoxynucleotides with novel modifications. *Nucleic Acids Res.* **12**, 7435–7453 (1984).
- Takeshita, M., Chang, C., Johnson, F., Will, S. & Grollman, A. P. Oligodeoxy-nucleotides containing synthetic abasic sites. Model substrates for DNA polymerases and apurinic/aprimidinic endonucleases. *J. Biol. Chem.* **262**, 10171–10179 (1987).
- Randall, S. K., Eritja, R., Kaplan, B. E., Petruska, J. & Goodman, M. F. Nucleotide insertion kinetics opposite abasic lesions in DNA. *J. Biol. Chem.* **262**, 6864–6870 (1987).
- Sagher, D. & Strauss, B. Insertion of nucleotides opposite apurinic/aprimidinic sites in deoxyribonucleic acid during *in vitro* synthesis: uniqueness of adenine nucleotides. *Biochemistry* **22**, 4518–4526 (1983).
- Schaaper, R. M., Kunkel, T. A. & Loeb, L. A. Infidelity of DNA synthesis associated with bypass of apurinic sites. *Proc. Natl Acad. Sci. USA* **80**, 487–491 (1983).
- Lawrence, C. W., Borden, A., Banerjee, S. K. & LeClerc, J. E. Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector. *Nucleic Acids Res.* **18**, 2153–2157 (1990).
- Paz-Elizur, T., Takeshita, M. & Livneh, Z. Mechanism of bypass synthesis through an abasic site analog by DNA polymerase I. *Biochemistry* **36**, 1766–1773 (1997).
- Doublie, S., Tabor, S., Long, A. M., Richardson, C. C. & Ellenberger, T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**, 251–258 (1998).
- Kiefer, J. R., Mao, C., Braman, J. C. & Beese, L. S. Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* **391**, 304–307 (1998).
- Huang, H., Chopra, R., Verdine, G. L. & Harrison, S. C. Structure of a covalently trapped catalytic

- complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**, 1669–1675 (1998).
22. Diederichsen, U. Selectivity of DNA replication: the importance of geometry over hydrogen bonding. *Angew. Chem.* **37**, 1655–1657 (1998).
23. Fyngenson, D. K. & Goodman, M. F. Appendix. Gel kinetic analysis of polymerase fidelity in the presence of multiple enzyme DNA encounters. *J. Biol. Chem.* **272**, 27931–27935 (1997).
24. Shibutani, S., Takeshita, M. & Grollman, A. P. Translesional synthesis on DNA templates containing a single abasic site. A mechanistic study of the “A rule”. *J. Biol. Chem.* **272**, 13916–13921 (1997).
25. Lindahl, T. & Nyberg, B. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11**, 3610–3618 (1972).
26. Weiss, B. & Grossman, L. Phosphodiesterases involved in DNA repair. *Adv. Enzymol. Rel. Areas Mol. Biol.* **60**, 1–34 (1972).
27. Ide, H. *et al.* Synthesis and damage specificity of a novel probe for the detection of abasic sites in DNA. *Biochemistry* **32**, 8276–8283 (1993).
28. Maulik, G. *et al.* Novel non-isotopic detection of MutY enzyme-recognized mismatches in DNA via ultrasensitive detection of aldehydes. *Nucleic Acids Res.* **27**, 1316–1322 (1999).
29. Denissenko, M. F., Pao, A., Tang, M. & Pfeifer, G. P. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in p53. *Science* **274**, 430–432 (1996).
30. Paris, P. L., Langenhan, J. & Kool, E. T. Probing DNA sequences in solution with a monomer–excimer fluorescence color change. *Nucleic Acids Res.* **26**, 3789–3793 (1998).

Acknowledgements. This work was supported by grants from the NIH and from the US Army Research Office.

Correspondence and requests for materials should be addressed to E.T.K. (e-mail: kool@chem.rochester.edu).

Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins

Uta-Maria Ohndorf*, Mark A. Rould†‡, Qing He*, Carl O. Pabo† & Stephen J. Lippard*

* Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Department of Biology and Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The anticancer activity of *cis*-diamminedichloroplatinum(II) (cisplatin) arises from its ability to damage DNA, with the major adducts formed being intrastrand d(GpG) and d(ApG) crosslinks¹. These crosslinks bend and unwind the duplex, and the altered structure attracts high-mobility-group domain (HMG) and other proteins². This binding of HMG-domain proteins to cisplatin-modified DNA has been postulated to mediate the antitumour properties of the drug^{3,4}. Many HMG-domain proteins recognize altered DNA structures such as four-way junctions and cisplatin-modified DNA⁵, but until now the molecular basis for this recognition was unknown. Here we describe mutagenesis, hydroxyl-radical footprinting and X-ray studies that elucidate the structure of a 1:1 cisplatin-modified DNA/HMG-domain complex. Domain A of the structure-specific HMG-domain protein HMG1 binds to the widened minor groove of a 16-base-pair DNA duplex containing a site-specific *cis*-[Pt(NH₃)₂][d(GpG)-N7(1),-N7(2)] adduct. The DNA is strongly kinked at a hydrophobic notch created at the platinum–DNA crosslink and protein binding extends exclusively to the 3′ side of the platinated strand. A phenylalanine residue at position 37 intercalates into a hydrophobic notch created at the platinum crosslinked d(GpG) site and binding of the domain is dramatically reduced in a mutant in which alanine is substituted for phenylalanine at this position.

HMG1 is a chromatin architectural factor having two tandem DNA-binding domains designated A and B (ref. 6). Both domains specifically recognize the 1,2-d(GpG) and 1,2-d(ApG) DNA intra-strand crosslinks formed by cisplatin^{7,8}. To investigate such a complex, a 16-base-pair double-stranded deoxyoligonucleotide containing a single *cis*-[Pt(NH₃)₂][d(GpG)-N7(G₈),-N7(G₉)] intrastrand crosslink (Fig. 1a) was crystallized as a 1:1 complex

with the 89-amino-acid domain A of HMG1 from rat (Fig. 1b). The 2.5 Å X-ray crystal structure of this protein–drug–DNA ternary complex was solved and refined (Fig. 1c, d, and Table 1), affording the first detailed picture, to our knowledge, of any complex between a structure-specific HMG domain and a DNA substrate.

As indicated in Fig. 1d, and as expected from previous studies of HMG domain–DNA complexes^{6,9,10}, the HMG domain binds with its concave surface contacting the minor groove of the DNA duplex. There is a sharp bend at the site of the cisplatin-induced DNA crosslink, but the relationship between this bend and the protein-binding site differs from that in the structures of sequence-specific HMG domains^{9,10} (Fig. 2a). These proteins induce a similar bend in DNA upon binding, but the bend locus is near the centre of the protein-binding site. The register of our HMG1 domain A with respect to the bend differs by two base pairs. Thus, the binding surface of HMG domain A extends exclusively to the 3′ side of the bend, when the platinated strand is used as a reference, and it covers five base-pair steps. In the SRY (ref. 9) and LEF-1 (ref. 10) complexes, residues near the start of helix I partially intercalate into the principal kink site. In HMG1 domain A, however, the aromatic side chain of Phe37 at the amino terminus of helix II intercalates into the hydrophobic cleft at the G₈–G₉ base-pair step formed by cisplatin crosslinking (Fig. 1d).

Hydroxyl-radical footprinting analysis confirms that this crystal structure represents the major binding mode that occurs in solution, as the protection pattern agrees with the protein–DNA contacts revealed in the crystal (Fig. 2b). The footprint covers four base pairs and is offset to the 3′ side of the platinum–d(GpG) crosslink. We also find that the ability of HMG1 domain A to bind to the cisplatin-modified oligonucleotide is significantly diminished in a Phe37Ala mutant, as revealed by gel mobility-shift analysis (Fig. 2c).

The precise positioning of the protein is directed not only by Phe 37 intercalation, but also by multiple interactions between DNA and amino-acid residues located on the N-terminal strand and in helices I and II. These contacts are shown schematically in Fig. 3a (see also Fig. 1a). Helix III, which has the largest number of charged residues, is primarily exposed to solvent and has only one, water-mediated, DNA contact. Helix I interacts with the sugar–phosphate backbone of the unmodified strand, whereas helix II mainly contacts the backbone of the platinated strand. The N termini of both helices fit in the minor groove. The presence of a hydrogen bond involving Ser 41 to A₁₀ could in part explain the higher affinity for cisplatin-modified DNA of HMG1 domain A compared to domain B, which has an alanine in that position⁷ (Fig. 1b).

The DNA in the complex is bent towards the major groove and the overall curvature¹¹ is 61° (Fig. 1d). Over the central six base pairs C₆–C₁₁, the minor groove is expanded to resemble A-type DNA, but is even more shallow. These features were previously encountered in cisplatin-modified DNA alone^{12–14}, from which it was suggested

Table 1 Data collection and refinement statistics

Crystal data	Native	I-dU-3	I-dU-5
Resolution (Å)	2.9	2.5	2.9
Unique reflections	6,389	9,655	6,238
R_{merge}^* (%)	6.0	4.4	9.3
R_{merge}^* (%) (last shell)	29.2	33.4	41.0
Completeness (%)	96.6	97.1	97.3
Completeness (%) (last shell)	97.7	84.2	96.0
Redundancy	2.9	3.7	2.9
Model refinement			
Resolution 100–2.5 Å	$R_{\text{cryst}} = 0.238^{\dagger}$	$R_{\text{free}} = 0.239^{\dagger}$	
	Protein	DNA	
R.m.s.d. bond length (Å)	0.014	0.011	
R.m.s.d. bond angle (°)	1.646	1.822	
R.m.s.d. planar dihedral angle (°)	1.451	1.664	

* $R_{\text{merge}} = \sum |I_i - \langle I \rangle| / \sum I_i$, where $\langle I \rangle$ is the average intensity over equivalent reflections.

† R_{cryst} and $R_{\text{free}} = \sum |F_o| - |F_c| / \sum |F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes. R_{free} was calculated with 10% of the reflections set aside during the refinement.

‡ Present address: Department of Molecular Physiology and Biophysics, University of Vermont College of Medicine, Burlington, Vermont 05405, USA.

that the ability of HMG-domain proteins to bind preferentially to distorted DNA structures might generally be explained by formation of a pre-bent duplex, with an accompanying widening of the minor groove in the DNA¹⁴. The propeller twist involving individual base pairs is much lower than in the platinated duplexes, and there is considerable buckling of the G₈/C₂₅ base pair (Fig. 3b).

The tertiary structure of domain A in the complex adopts the usual L-shaped fold comprising three α -helices connected by two loops (Fig. 1d). Several clusters of highly conserved hydrophobic residues stabilize interactions of helices I and II, and interactions between helix III and the N-terminal strand. The structures of bound HMG1 domain A reported here, and that of the free domain in solution¹⁵, superimpose with a r.m.s.d. of 2.1 Å. The conformational changes induced in the protein upon DNA binding occur primarily in two areas. First, the region encompassing the C terminus of helix I, the N terminus of helix II, and the connecting loop, adjust to allow for intercalation of Phe 37. In addition, the N-terminal strand and helix III, which together form the longer arm of the L, are slightly displaced compared with their positions in the NMR structure in order to optimize interactions with the sugar-phosphate backbone of the DNA. In agreement with measurements

of backbone dynamics for uncomplexed HMG1 domain A in solution¹⁶, loop 1 (residues 28–37) is quite flexible, even after binding to the DNA substrate. This flexibility is revealed by larger temperature factors for the backbone atoms, as well as by the relatively weak electron density for the side chains of residues 28–30, 32 and 34.

The *cis*-diammineplatinum(II) fragment is anchored to the N7 atoms of purine rings G₈ and G₉. The Pt–N7 distances were not restrained, yet refined to the expected value¹⁷ of 2.0 ± 0.1 Å. One of the NH₃ ligands is within 3.2 Å one of the phosphate oxygen atom of G₈, which is close enough to participate in hydrogen bonding, as found previously in the structure of cisplatin-modified d(pGpG)¹⁷. The large roll¹⁸ of 61° induced by the platinum crosslink destacks the G₈–G₉ purine rings and exposes their hydrophobic surfaces in the minor groove. The phenyl ring of Phe 37 serves as a wedge, projecting into this hydrophobic crevice and stacking onto the G₉ base at a distance of 3.5 Å (Fig. 3b). The dihedral angle between the two planes of the cisplatin-coordinated guanine heterocycles is 75° (Fig. 3c). This angle is considerably larger than the dihedral angles observed in the X-ray and NMR structures of duplexes containing the 1,2-intrastrand d(GpG) crosslink^{12–14,19}, which range from 30° to

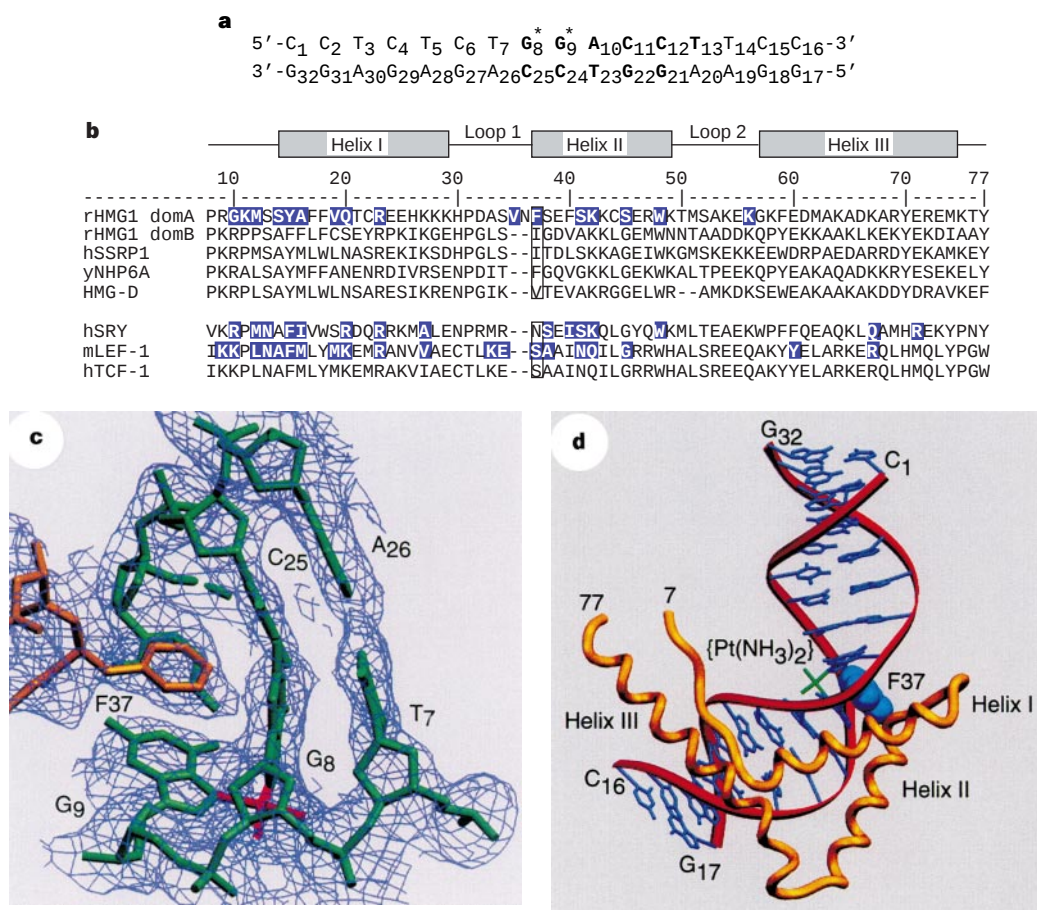


Figure 1 Primary structure and features of the complex between the non-sequence-specific domain A of HMG1 and cisplatin-modified DNA. **a**, Sequence of DNA used here; asterisks, nucleotides crosslinked by cisplatin. Nucleotides contacted by the protein are in bold type. **b**, Alignment of selected members of the two subfamilies of HMG-domain proteins³⁰, based on primary sequence and structural superposition. Shaded boxes represent α -helices and solid lines indicate loops. The top group are a representative set of structure-specific HMG domains and the bottom group are sequence-specific proteins. The HMG1 domain A sequence and numbering shown correspond to the residues located and refined in the crystal structure. Residues that contact DNA in structurally characterized complexes are highlighted. Amino acids at positions

equivalent to residue 37 of HMG1 are boxed. **c**, A simulated annealing omit map (starting temperature, 2,000 K), contoured at the 1.1 σ level, in the region of the cisplatin–DNA adduct and showing the localized bend at the platination site and the intercalated Phe 37 side chain. For computing this ($F_o - F_c$) electron density map, residues 36–38 and bases T₇–G₉, C₂₄–A₂₆ were omitted. **d**, Overall structure of the complex. The protein backbone is shown in yellow, the intercalating Phe 37 residue as van der Waals spheres, and the DNA in red and blue with the *cis*-[Pt(NH₃)₂{d(GpG)-N7(G₈)-N7(G₉)}] intrastrand adduct in green. Numbers indicate the first (N terminus) and last (C terminus) ordered residues in the crystal structure. Main-chain torsion angles for all non-glycine residues fall within energetically favourable Ramachandran boundaries.

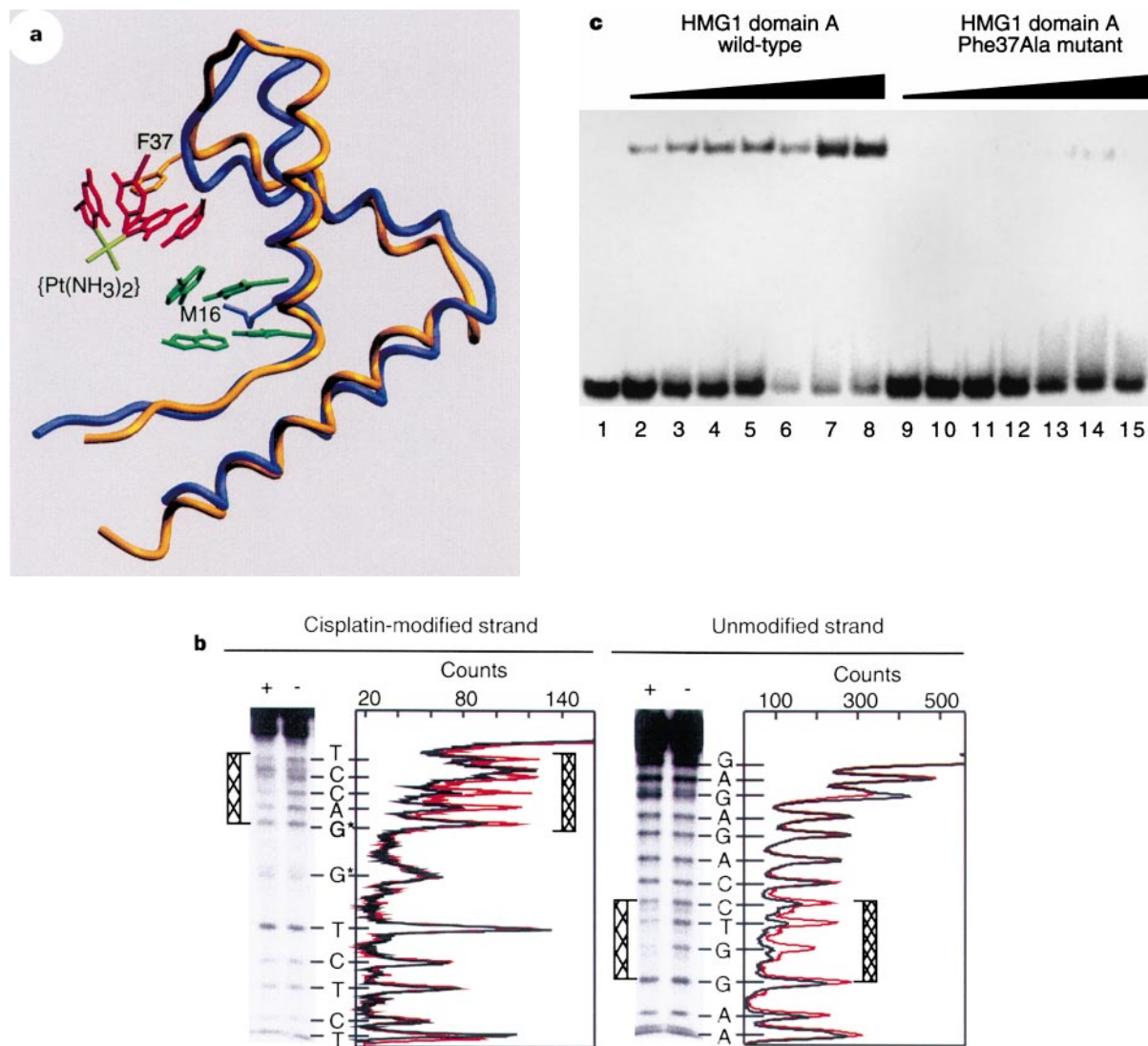


Figure 2 Structure of the HMG1 domain A complex with cisplatin-modified DNA. **a**, Superposition of the protein components of a sequence-specific LEF-1 DNA complex¹⁰ (PDB entry 2LEF) in blue, with the present structure in yellow, illustrates the difference in docking modes. Two base pairs constituting the primary bend locus in the LEF-1 DNA complex are shown in green, with the intercalating Met 16 residue in blue. This bend is offset by two base pairs from the bend locus at the cisplatin-crosslinked G/C base pairs in the present complex, depicted in red. **b**, Hydroxyl-radical footprint of HMG1 domain A bound to cisplatin-modified DNA in the presence (+, black trace) or absence (–, red trace) of HMG1 domain. A

footprint extending over 3–4 bases on both strands, as indicated by the hatched bars, is revealed in the autoradiograph as well as in the corresponding scaled phosphorimager traces. **c**, Gel-mobility shift assays comparing the DNA-binding ability of wild-type HMG1 domain A (lanes 2–8) with that of a Phe37Ala mutant (lanes 9–15). Lane 1 contains no protein. The cisplatin-modified probe (5 nM) was incubated with 10, 20, 40, 60, 80, 100 and 200 nM protein. The band at higher mobility corresponds to unbound DNA, whereas that of lower mobility represents the specific protein–DNA complex.

60°, and is more like that in the *cis*-[Pt(NH₃)₂{d(pGpG)}] structure (Fig. 3c). The resulting more favourable platinum coordination geometry, an effect that lowers the free energy of complex formation, is further stabilized by the stacking interactions between Phe 37 and G₉ (Fig. 3b).

The conformational changes induced in linear DNA by the sequence-specific HMG domains of SRY (ref. 9) or LEF-1 (ref. 10) are similar to those found in the present complex. Despite the different positioning of the domains with respect to the principal DNA bend, the sugar–phosphate backbones in the sequence- and structure-specific HMG-domain–DNA complexes superimpose well on one another. The novel DNA-binding motif adopted by the HMG1 domain observed here may represent an alternative binding mode that is available to structure-specific HMG domains^{7,20–22}. Examination of sequence alignments reveals a striking difference in the nature of the residue at position 37 for structure- and sequence-specific HMG domains (Fig. 1b). Nearly

all chromatin-associated HMG-domain proteins, including HMG1/2, HMG-D and NHP6A, have a hydrophobic residue at this position that can serve as a bending wedge, but which does not seem to discriminate on the basis of sequence. The HMG domains found in transcription factors such as LEF-1, TCF-1 and SRY, however, have a polar residue at position 37, which would allow sequence-specific hydrogen-bond formation.

A molecular dynamics study of the structure-specific HMG-D protein with a DNA substrate pre-bent by a disulphide crosslink predicted the intercalation of both Met 16 and Val 37 (ref. 23) (Fig. 1b). The methyl side chain of Ala 16 in HMG1 domain A is too short to intercalate into the G₂₂–T₂₃ base step, however, and instead forms a hydrophobic contact with the ribose ring of T₂₃. Most other sequence-neutral HMG domains have larger hydrophobic residues at position 16, which could intercalate in the same way as sequence-specific HMG domains. The binding mode seen here could thus be adopted by a subset of structure-specific HMG-domain proteins,

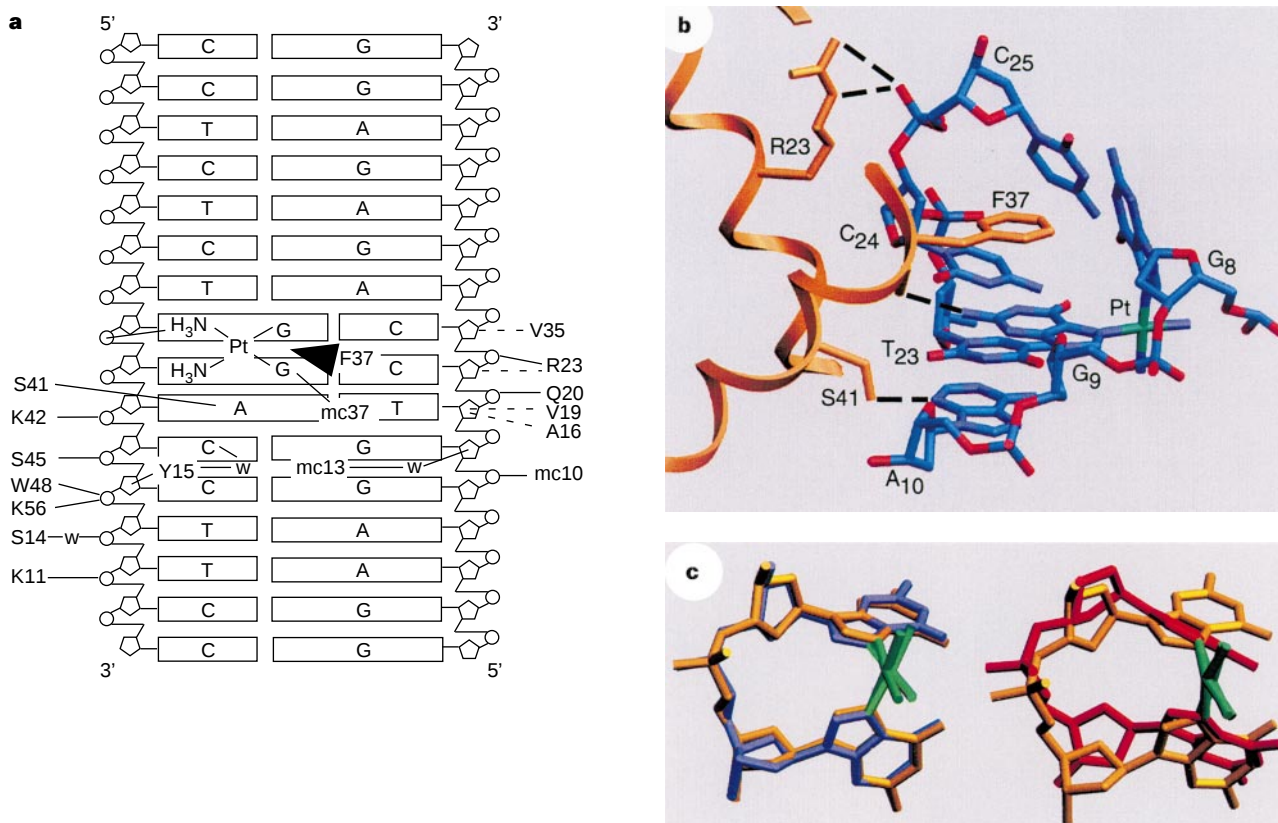


Figure 3 The platinum-DNA crosslink and protein-DNA interactions. **a**, Diagram summarizing all HMG domain-DNA contacts. Solid lines, hydrogen bonds or salt bridges; dashed lines, van der Waals contacts; mc, main chain; and w, water molecule. The wedge indicates the Phe-37 intercalation site. **b**, Expanded view of protein-DNA interactions at the site of platination. The aromatic rings of Phe37 and G₈ pack edge-to-face with a dihedral angle of 74°. The main-chain carbonyl group of Phe37 is within hydrogen-bonding distance (2.9 Å) of the exocyclic amino group of G₈. As predicted⁷, Ser41 is involved in a very tight

hydrogen-bonding contact with N3 of A₁₀. **c**, Superposition of crystallographically determined cisplatin-[d(GpG)] intrastrand crosslinks in the present structure (yellow), the equivalent site in the structure of a single-stranded stranded *cis*-[Pt(NH₃)₂(d(pGpG))] platinated dinucleotide¹⁷ (blue, left), and the site in a cisplatin-modified DNA dodecamer duplex¹⁴ (red, right), portraying dihedral angles between the two coordinated guanine bases of ~75°, ~77° and ~30°, respectively, as calculated by the program QUANTA 4.1 (Molecular Simulations).

which would account for the differences in DNA-structure recognition and activity of domains A and B of HMG2 (ref. 24).

The asymmetric placement of the protein with respect to the cisplatin adduct suggests that another HMG domain or cellular protein might be accommodated at the 5' side of the crosslink. In our crystals, helix II of an adjacent HMG-domain-DNA complex contacts this site (to be described elsewhere), although our footprinting result indicates that this additional interaction does not occur in solution.

An understanding of how structure-specific HMG-domain proteins bind to cisplatin-modified DNA should provide a rational basis for the design of new chemotherapeutic strategies and of molecules with an increased affinity and specificity for cisplatin-modified DNA and for new platinum complexes. Such agents could enhance the cytotoxic effect of cisplatin and related compounds by shielding damaged DNA from intracellular repair.

Note added in proof: After submission of this paper, a model structure, constructed based on NMR data of a DNA complex of NHP6A, was reported by F. H.-T. Allain *et al.* (*EMBO J.* **18**, 2563–2579; 1999). □

Methods

Expression and purification. Recombinant domain A (residues M1-F89)²⁵ from HMG1 was prepared and purified as described⁷, but with an additional heparin affinity-column purification step.

Hydroxyl-radical footprint. The 16-base-pair oligonucleotide was ³²P-labelled

at its 5' end either on the platinum-crosslinked or on the unmodified strand and subjected to Fe(EDTA)-mediated hydroxyl-radical cleavage for 13 min in the presence or absence of a 20-fold excess of HMG1 domain A over DNA. The cleavage reaction was stopped by addition of thiourea.

Crystallization, structural determination and refinement. Crystals of HMG1 domain A bound to a cisplatin-modified 16-base-pair deoxyoligonucleotide were obtained by vapour diffusion. Solutions containing 0.6 mM protein, 0.9 mM DNA duplex, 0.05 M HEPES, pH 6.4, 40 mM magnesium acetate, 1% glycerol, 8% PEG 3350, and 2.5 mM DTT were equilibrated in hanging drops over solutions containing 0.1 M HEPES, pH 6.3, 80 mM magnesium acetate, 1% glycerol, 16% PEG 3350, and 5 mM DTT for 10 days at 20°. Crystals formed in space group C2 with *a* = 105.21 Å, *b* = 50.41 Å, *c* = 53.13 Å, β = 99.82°, and one complex per asymmetric unit. Diffraction data were collected at -150 °C on a Rigaku R-Axis IIC Image-plate area detector system, processed with DENZO/SCALEPACK²⁶ and scaled by using local scaling (MAXSCALE)²⁷. Multiple isomorphous replacement phases, including anomalous data (MIRAS), were calculated from two iodinated derivatives, denoted as I-dU3 and I-dU5, having 5-iodouracil incorporated at positions T₃ and T₅ (Fig. 1a), respectively. Heavy-atom coordinates were determined from a difference Patterson map and refined with the program MLPHARE²⁸. The initial MIRAS electron-density map was improved by solvent flattening with the program DM²⁸. Canonical form-B DNA was built into the density, a polyalanine model of domain A was constructed, and side chains for all residues with clear density were added to the model. Disordered side chains of residues 7, 9, 28–30, 32, 34 and 64 were not located. The model was built and manually adjusted with the program vuSette zc (M.A.R., unpublished results)

and refined with XPLOR²⁹. The model was refined initially against data collected from the native crystal. Subsequent phase extension to 2.5 Å was carried out with data from the I–dU3 derivative.

Received 29 December 1998; accepted 12 May 1999.

- Lepre, C. A. & Lippard, S. J. in *Nucleic Acids and Molecular Biology* (eds Eckstein, F. & Lilley, D. M. J.) 9–38 (Springer, Heidelberg, 1990).
- Zlatanova, J., Yaneva, J. & Leuba, S. H. Proteins that specifically recognize cisplatin-damaged DNA: a clue to anticancer activity of cisplatin. *FASEB J.* **12**, 791–799 (1998).
- Zamble, D. B. & Lippard, S. J. Cisplatin and DNA repair in cancer chemotherapy. *Trends Biochem. Sci.* **20**, 435–439 (1995).
- Zhai, X., Beckmann, H., Jantzen, H.-M. & Essigmann, J. M. Cisplatin–DNA adducts inhibit ribosomal RNA synthesis by hijacking the transcription factor human upstream binding factor. *Biochemistry* **37**, 16307–16315 (1998).
- Bustin, M. & Reeves, R. in *Progr. Nucleic Acid Res. Mol. Biol.* (eds Cohn, W. E. & Moldave, K.) 35–100 (Academic, San Diego, 1996).
- Read, C. M. *et al.* in *Nucleic Acids and Molecular Biology* (eds Eckstein, F. & Lilley, D. M. J.) 222–250 (Springer, Berlin, 1995).
- Dunham, S. U. & Lippard, S. J. DNA sequence context and protein composition modulate HMG-domain protein recognition of cisplatin-modified DNA. *Biochemistry* **36**, 11428–11436 (1997).
- Chow, C. S., Barnes, C. M. & Lippard, S. J. A single HMG domain in high-mobility group 1 protein binds to DNAs as small as 20 base pairs containing the major cisplatin adduct. *Biochemistry* **34**, 2956–2964 (1995).
- Werner, M. H., Huth, J. R., Gronenborn, A. M. & Clore, G. M. Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY–DNA complex. *Cell* **81**, 705–714 (1995).
- Love, J. J. *et al.* Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* **376**, 791–795 (1995).
- Lavery, R. & Sklenar, H. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **6**, 63–91 (1988).
- Dunham, S. U., Dunham, S. U., Turner, C. J. & Lippard, S. J. Solution structure of a DNA duplex containing a nitroxide spin-labeled platinum d(GpG) intrastrand cross-link refined with NMR-derived long-range electron–proton distance restraints. *J. Am. Chem. Soc.* **120**, 5395–5406 (1998).
- Gelasco, A. & Lippard, S. J. NMR solution structure of a DNA dodecamer duplex containing a *cis*-diammineplatinum(II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry* **37**, 9230–9239 (1998).
- Takahara, P. M., Frederick, C. A. & Lippard, S. J. Crystal structure of the anticancer drug cisplatin bound to duplex DNA. *J. Am. Chem. Soc.* **118**, 12309–12321 (1996).
- Hardman, C. H. *et al.* Structure of the A-domain of HMG1 and its interaction with DNA as studied by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry* **34**, 16596–16607 (1995).
- Broadhurst, R. W., Hardman, C. H., Thomas, J. O. & Laue, E. D. Backbone dynamics of the A-domain of HMG1 as studied by ¹⁵N NMR spectroscopy. *Biochemistry* **34**, 16608–16617 (1995).
- Sherman, S. E., Gibson, D., Wang, A. H.-J. & Lippard, S. J. Crystal and molecular structure of *cis*-[Pt(NH₃)₂{d(pGpG)}], the principal adduct formed by *cis*-diamminedichloroplatinum(II) with DNA. *J. Am. Chem. Soc.* **110**, 7368–7381 (1988).
- Dickerson, R. E. DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.* **26**, 1906–1926 (1998).
- Yang, D., van Boom, S. S. G. E., Reedijk, J., van Boom, J. H. & Wang, A. H.-J. Structure and isomerization of an intrastrand cisplatin–crosslinked octamer DNA duplex by NMR analysis. *Biochemistry* **34**, 12912–12920 (1995).
- Teo, S.-H., Grasser, K. D. & Thomas, J. O. Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY. *Eur. J. Biochem.* **230**, 943–950 (1995).
- Weiss, M. A., Ukiyama, E. & King, C.-Y. The SRY cantilever motif discriminates between sequence- and structure-specific DNA recognition: alanine mutagenesis of an HMG box. *J. Biomol. Struct. Dyn.* **15**, 177–184 (1997).
- Berners-Price, S. J. *et al.* Structural transitions of a GG-platinated DNA duplex induced by pH, temperature and box A of high-mobility group protein 1. *Eur. J. Biochem.* **243**, 782–791 (1997).
- Balaeff, A., Churchill, M. E. A. & Schulten, K. Structure prediction of a complex between the chromosomal protein HMG-D and DNA. *Proteins* **30**, 113–135 (1998).
- Yoshioka, K. *et al.* Differences in DNA recognition and conformational change activity between boxes A and B in HMG2 protein. *Biochemistry* **38**, 589–595 (1999).
- Falciola, L., Murchie, A. I. H., Lilley, D. M. J. & Bianchi, M. E. Mutational analysis of the DNA binding domain A of chromosomal protein HMG1. *Nucleic Acids Res.* **22**, 285–292 (1994).
- Otwiński, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Rould, M. A. Screening for heavy-atom derivatives and obtaining accurate isomorphous differences. *Methods Enzymol.* **276**, 461–472 (1997).
- Collaborative Computational Project number 4. The CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Brünger, A. T. XPLOR Version 3.1, A system for X-ray crystallography and NMR (Yale University Press, New Haven, Connecticut, 1992).
- Baxeianis, A. D. & Landsman, D. The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res.* **23**, 1604–1613 (1995).

Acknowledgements. This work was supported by a grant from the National Cancer Institute (to S.J.L.), by the Howard Hughes Medical Institute (C.O.P.), by equipment purchased with support from the PEW Charitable Trusts, and by predoctoral fellowships from the Fonds der Chemischen Industrie, Germany (U.-M.O.), and the Howard Hughes Medical Institute (Q.H.). We thank S. U. Dunham for discussion, and J. S. Wishnok for mass spectrometric analysis.

Correspondence and requests for materials should be addressed to S.J.L. (e-mail: lippard@lippard.mit.edu). The atomic coordinates have been deposited in the Brookhaven Protein Database (accession number, 1ckt).

Looking for a solution

New biochemicals, the latest catalogues and (where would reagents be without it?) laboratory quality water constitute this 'biochemicals and reagents' collection.

Acids

From Mallinckrodt Baker www.jtbaker.nl

High-grade acid preparations

The specification of the J. T. Baker range of acids has been improved to reduce the level of trace metal impurities. Baker Analyzed acids are specified for at least 25 trace metal impurities at sub-p.p.m. level. A special low-mercury grade (less than 5 p.p.b.) is available for sulphuric, nitric and hydrochloric acids.

Reader Service No. 100



Acid test: minimizing the metals.

Organics web site

From Acros Organics www.be.acros.com

A not-so-loony 'toon'

The new Acros Organics website includes a structure search facility which works without the need to download any software. The site features cartoon characters drawn by a "renowned Belgian cartoonist". The site contains information on the company's range of 16,000 organic chemicals. Searching is by product name, partial name, CAS number or formula. The structure search simply requires the structure to be drawn on the screen.

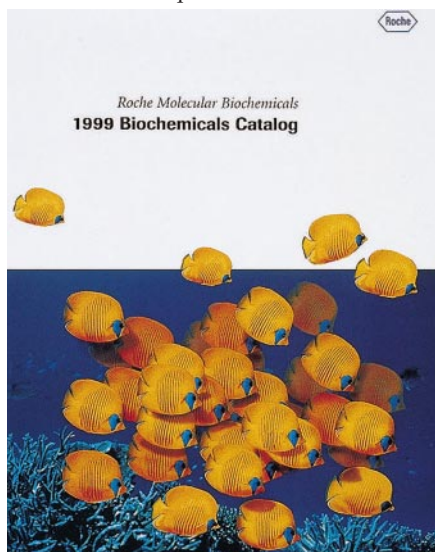
Reader Service No. 101

Biochemicals

From Roche www.roche.com/diagnostics

Roche's 1999 Biochemicals Catalog

This new catalogue covers molecular biology, cell biology, immunochemistry, biochemistry, laboratory reagents and products for drug and disease research. The new instrument section gives details on the LightCycler instrument for quantitative PCR and the



Gone fishing: tackle for the biochemist.

Ecdysteroids

From Alexis Biochemicals www.alexis-corp.com

Large or small quantities of ecdysteroids

Alexis Biochemicals offers a wide range of rare and highly purified ecdysteroids, tetracyclic polyhydroxylated compounds present in exceedingly small amounts in most classes of invertebrates. The range includes muristerone A, ponasterone A, ecdysone, hydroxyecdysone and makisterone A.

Reader Service No. 103

Aqualine

From Fisher Scientific www.fisher.co.uk

Extending the range

The Aqualine range of pyridine-free Karl Fischer reagents has now been expanded to include coulometric reagents. Coulometric reagents are normally used in preference to volumetric reagents where analysed water levels are much lower. The new coulometric reagents comprise Electrolyte A, a cathodic reagent, and an anodic reagent, Electrolyte C.

Reader Service No. 104

FP syringe filter holders

From Schleicher & Schuell www.s-and-s.com

Ready-to-use filter holders

The FP range of syringe filter holders incorporates a large filtration area of 5.7 cm² for fast filtration and very low 'dead' volumes. These filter holders are available in a wide choice of pore sizes and diameters with different connectors. The filters are suitable for cold sterilization, particle filtration, clarification or pre-filtration of volumes from 1 ml up to 1,000 ml. Applications include tissue culture media, serum, biological fluids,



Filter tips: FP from S&S.

reagents, buffer and aqueous solutions in HPLC. Pore sizes range from 0.2 to 5.0 µm.

Reader Service No. 105

Chemicals

From Pierce www.piercenet.com

This year's catalogue

Pierce Chemical Company's latest catalogue features immunology, protein chemistry, molecular biology, high-throughput screening, high-performance dialysis and gas chromatography products. For a copy call 800-874-3723 or 815968-0747 in the USA. And the website includes a literature order form.

Reader Service No. 106

Vision

From Atas www.atas-uk.com

Sample preparation for water analysis

The Vision Sample Preparation System was developed for automatic online determination of trace organics in water. Coupled to a capillary GC/MS, it provides 24-hour operation for routine analysis of drinking water and environmental samples. Sensitivity is further enhanced by large-volume injection using the Optic 2 Programmed Injector.

Reader Service No. 107

Milli-Q

From Millipore www.millipore.com

Water purification to p.p.t. levels

The Milli-Q Element Water System purifies water down to p.p.t. levels for critical laboratory applications. The system incorporates UV photo-oxidation to destroy organics and allows for foot-pedal operation to minimize cross-contamination by users and for operation under a laminar hood. The Milli-Q Element System is designed for laboratories that need ultrapure water for blank, standard and sample preparations for critical trace analyses at ng/l and pg/l levels.

Reader Service No. 108

Water Purification

From USF Elga

www.elga.co.uk

Two ranges merge

USF Elga, formed by the merger of USF and Elga, has launched a new range of water purification systems. The new range unites USF and Elga products behind a new colour scheme, and the systems are supported by a network of over 500 service centres.

Reader Service No. 109

Label-IT

From Mirus

www.genetransfer.com

Cy3 and Cy5 labelling kept simple

Mirus Corporation's Label-IT Non-Radioactive Labelling Reagents have been developed for single-step labelling with Cy3 and Cy5 dyes. Kits are now available for labelling DNA and RNA. The protocol eliminates the need for enzymatic labelling reactions, which use large numbers of complex reagents.

Reader Service No. 110

TransFast

From Promega

www.promega.com

A serum-free transfection reagent

The TransFast transfection reagent was developed to perform well in many cell lines, giving quick results for transient and stable



Flying new colours: Elga goes white and blue.

transfection. Each kit includes transfection reagent, nuclease-free water and protocol.

Reader Service No. 111

UV-visible spectroscopy

From Hewlett-Packard

www.hp.com/go/chem

Bilingual spectroscopy

Hewlett-Packard a présenté une version française du système de spectroscopie du HP 845E visible-UV. Les utilisateurs ont deux options pour la mise en service du logiciel: anglais ou français. Une fois choisi et relancé, le système affiche le langage voulu.

Numéro 112 De Service De Lecteur

Peroxidase substrates

From InnoGenex

www.innogenex.com

Single-reagent peroxidase substrates

AEC (3-amino-9-ethyl-carbazole) and DAB (3,3'-diaminobenzidine tetrahydrochloride)

are peroxidase substrates that produce coloured precipitates on enzymatic conversion. They are used for calorimetric detection of antibodies and probes labelled directly or indirectly with horseradish peroxidase. InnoGenex now produce AEC and DAB as single-reagent formulations, offering the advantage of a ready-to-use solution with long-term stability.

Reader Service No. 113

1999/2000 Catalogue

From Vector Laboratories www.vectorlabs.com

A catalogue from specialists in nonradioactive detection systems and lectin production

Highlights in the 1999/2000 edition include: the Mouse on Mouse (M.O.M.) immunodetection kit for using mouse monoclonals on mouse tissues, especially useful in transgenic, knock-out and xenograph murine models; and the Vector NovaRED peroxidase substrate kit, which offers a sensitive, brick-red-coloured precipitate that is insoluble in ethanols and organic solvents like xylene and can be permanently mounted. It is ideal for immunohistochemistry and blotting.

Reader Service No. 114

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card.

ADVERTISEMENTS

STERNBERGER MONOCLONAL (SMI) ANTIBODIES FROM AFFINITI

The SMI series of antibodies to neurofilaments, tau, myelin basic protein, blood brain barrier (EBA), SNAP-25, and tubulins is available from AFFINITI.

We also offer many antibodies to proteins of neuroscientific interest, including GAD65/67, α , β and γ (NEW) Synucleins, heme oxygenase, neuronal nitric oxide synthase, PGP9.5 and β -catenin.

For data sheets or more information on AFFINITI's extensive range of products for neuroscience research please visit our website at:

[HTTP://WWW.AFFINITI-RES.COM](http://www.affiniti-res.com)

AFFINITI Research Products Ltd., Mamhead Castle, EXETER, EX6 8HD, UK
Tel. +44/0 1626 891010; Fax: +44/0 1626 891090; URL: <http://www.affiniti-res.com>

READER ENQUIRY NO. 15

Diagnostic Reagents

Pefachrome®-Series: chromogenic and fluorogenic substrates

Pefabloc®-Series:
serine proteinase inhibitors

Purified snake venom proteins
(e.g. Protac®, Ecarin, Batroxobin)

PENTAPHARM LTD

CH 4002 Basel Fax: +41 61 319 96 19
<http://www.pentapharm.ch>

READER ENQUIRY NO. 16

The Mouse Kit™ for Gene Targeting in ES Cells

- ✓ Complete with embryonic stem (ES) cells, feeder cells and essential reagents (including FBS) for engineering one knockout mouse strain
- ✓ Low-passage AB2.2-Prime ES Cells* (129/SvEv)
- ✓ Uniform, high-quality ESQ feeder cells (mitotically inactivated STO cells) engineered for *neo* and *puro* resistance
- ✓ Prequalification of ESQ feeder cells and reagents for use with AB2.2-Prime ES Cells

*Manufactured by Lexicon Genetics Incorporated

Fax: (619) 535-0071
Tel: (619) 535-5400 or (800) 424-5444
www.stratagene.com

READER ENQUIRY NO. 6

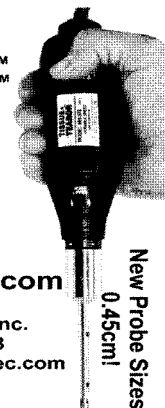
Homogenizers?

- Tissue-Tearor™
- BeadBeater™
- Mini-BeadBeater™
- BioHomogenizer™

Check
Us
Out

www.biospec.com

BioSpec Products, Inc.
Ph/Fax 918-336-3363
e-mail info@biospec.com



READER ENQUIRY NO. 54

Lysyl Endopeptidase (Achromobacter Protease 1)

Specifically cleaves all Lys-X bonds including Lys-Pro.
Highly Stable, Water Soluble

Wako

Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 5